The Dissertation Committee for Michael Charles Yates certifies that this is the approved version of the following dissertation:

NEW PERSPECTIVES ON THE DETERMINANTS AND CONSEQUENCES OF INDIVIDUALS' INVESTMENT DECISIONS

Committee:

____
Laura Starks, Supervisor

____
John Griffin

____
Alok Kumar

____
Robert Lieli

____
Paul Tetlock

NEW PERSPECTIVES ON THE DETERMINANTS AND CONSEQUENCES OF INDIVIDUALS' INVESTMENT DECISIONS

by

Michael Charles Yates, B.S.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2007

Dedication

For Kimberly

Acknowledgements

I would like to thank the members of my doctoral committee for their hard work and assistance. In particular, I recognize my dissertation chair, Laura Starks, for her valuable guidance throughout the job market and dissertation process. I also thank John Griffin, who has acted as a supervisor, a mentor and a friend to me for the past four years. Comments and advice from Paul Tetlock, Alok Kumar and Robert Lieli have greatly benefited my writing and my professional development.

I also recognize Johan Sulaeman, whose astounding unselfishness and willingness to help his fellow student have made him a dear friend and an invaluable resource to me over the past five years.

I thank my parents for setting my feet on the pathway to success. Their unconditional love, support, training and encouragement are imprinted upon my every accomplishment. I also thank countless other friends and family who have offered encouragement over the years.

I thank my wife Kimberly for her love and support, as well as her unquestioning acceptance of the student lifestyle we have shared since the day we married. Her eyes are a mirror through which my reflection always appears stronger and more capable than I truly am, and her unwavering belief in me has helped me through the stress and uncertainty of the past year more than I could ever express.

Finally, I acknowledge He who makes all things possible for a richness of blessings beyond what any man deserves that has culminated in the completion of this dissertation.

NEW PERSPECTIVES ON THE DETERMINANTS AND CONSEQUENCES OF INDIVIDUALS' INVESTMENT DECISIONS

Publication No. ___________

Michael Charles Yates, Ph.D.
The University of Texas at Austin, 2007

Supervisor: Laura Starks

This research examines how individuals formulate their investment decisions and the importance of these decisions to the financial marketplace. Traditional finance theory has focused on solving the rational investor's choice problem by considering each financial asset's contribution to the risk and return of the investor's existing portfolio. Alternatively, this study recognizes the inability of most individuals to consider all possible investments in the financial universe, and therefore approaches the investor's choice problem by focusing on environmental and psychological factors that guide the formulation of the investor's selection set. In particular, this research focuses on the importance of attention in influencing the common stock selections of individuals and shows that this attention effect can have a significant impact on the returns of attention-grabbing equities. Additionally, I document the impact of mutual fund family affiliation on the mutual fund investment decisions of individuals and discuss how apparent reputation effects could impact the organization and performance incentives of mutual funds.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In traditional finance theory, the investor's choice problem is relatively straightforward. After determining the optimal amount of consumption and investment, the unconstrained theoretical investor selects from all available investment vehicles in the universe so as to maximize his or her portfolio's expected return relative to its risk. In reality, however, investors are not truly unconstrained, and there exists a staggering number of real and financial investments from which an investor may choose. Limited time and bounded cognitive ability render it infeasible for a real-world investor to know and understand each potential investment's expected contribution to the risk and return of the investor's existing portfolio.

In contrast to traditional theory, this study recognizes the inability of most individuals to consider all possible investments in the financial universe, and therefore approaches the investor's choice problem by focusing on the environmental and psychological factors that guide the formulation of the investor's selection set. In particular, this research focuses on the importance of attention in influencing the common stock selections of individuals and shows that attention can have a significant impact on the returns of attention-grabbing equities. Additionally, I document the importance of mutual fund family affiliation on the mutual fund investment decisions of individuals and discuss how these effects might impact the organization and performance incentives of mutual funds.

In Chapter 3, I analyze the effect of attention on the stock choices of individual investors and the returns of stocks that attract such attention. I show that attention-grabbing events can predict not only the timing and direction of individual trades, but also stock price returns at short horizons. My research on the effects of attention on stock market trading and prices provides two major contributions to the literature. First, by focusing on attention related to uninformative events such as 52-week highs and lows, my paper is the first to divorce the effects of attention from underlying information and, thus, is the first to estimate the magnitude of pure attention effects in the financial marketplace. Second, by showing that behavioral influences such as limited attention can impact stock returns, I underscore the importance of studying and understanding the psychological factors that affect individuals' investment behavior.

My key findings relating to attention and its implications for the stock market are as follows. First, I find that stocks experiencing 52-week highs and lows earn statistically and economically significant abnormal returns on days immediately following the event. The effect remains positive and significant after controlling for various event-day, firm-level variables such as return, turnover, volatility and liquidity. Furthermore, the return predictability exists even for 52-week highs and lows with below average event-day absolute return, turnover or intraday volatility, further reinforcing the notion that underlying information is not driving my predictability results. I additionally find that days on which the attention grabbing event is most salient result in significantly larger next-day abnormal returns. In a final test of the whether attention truly causes my results, I show that similarly-defined highs and lows over shorter reference periods (e.g. a 51-

week high) that are not simultaneously 52-week highs or lows do not significantly affect next day returns. This result is particularly revealing, since the stocks experiencing these alternative high and low events should differ from 52-week high and low stocks only in the amount of attention they receive.

Turning next to the effect of attention on individuals' stock choices, I show that sample individual traders at a large discount brokerage have significantly higher net buy imbalances for stocks that experienced 52-week highs or lows on the previous day, indicating that attention plays an important role in determining the consideration sets of these sample investors. Finally, I show evidence that the abnormal positive returns predicted by 52-week highs are confined to those occasions when individuals are net buyers following the event, suggesting that individual trading may be the conduit through which these short-term abnormal returns are generated.

In Chapter 4, I study the role that fund family membership plays in determining the beliefs and ultimately the mutual fund holdings of individual investors. This research is the first to empirically study the effect that fund family reputation has on the mutual fund trades of individuals. The main findings presented in Chapter 4 are as follows. First, I show that fund family membership plays an important role for individual investors. Sample individuals at a large discount brokerage show a significant propensity to concentrate their mutual fund portfolios at the family level. This behavior is particularly noteworthy because the investors in my sample buy funds through omnibus accounts and thus do not receive direct service from the fund advisor. Furthermore, such accounts remove other institutional constraints that might explain why an investor would stay within a particular fund family such as for ease of switching funds, for lessening of search costs, or due to restrictions caused by employee retirement plans. Furthermore, I find that individuals' beliefs about fund families do not change swiftly over time, as sample investors are observed to select a disproportionate number of funds from families with which they have previous investment experience. Indeed, regression analysis reveals that a fund is at least six times more likely to be selected if it belongs to a family with which the purchasing investor has prior experience. Moreover, I show that this effect of previous family ownership is consistent for investors who experience positive or negative returns with the fund family, indicating that a major component of the beliefs of sample individuals in the competence of particular families arises separately from past performance.

I interpret this inclination by individuals to repeatedly purchase funds from the same family, even when they have experienced poor return performance with the family, as evidence that family reputation plays an important role in the mutual fund choice decision. Consistent with this explanation, but inconsistent with alternative explanations such as family style preferences or familiarity arising from shared prospectuses within family objective classes, I find that the magnitude and significance of the effect of previous family ownership on the probability that a family fund is selected in the future obtains even for investors choosing funds with investment objectives that are new to the investor. Moreover, I demonstrate that individuals' beliefs about funds belonging to older and larger families dissipate only gradually, as evidenced by decreased flow-performance sensitivity for these funds. To the extent that family size and age are reasonable proxies for a family's reputation, this finding is consistent with the hypothesis that investors' convictions in families with more established reputations change more slowly over time.

The remainder of this dissertation proceeds as follows. Chapter 2 reviews the relevant literature relating to the research conducted herein. Chapter 3 examines the effect of attention on the trading behavior of individuals and the returns of attention-grabbing stocks. Chapter 4 presents evidence on the importance of mutual fund families to the purchases and belief formation of individual investors. Chapter 5 concludes.

CHAPTER 2: LITERATURE REVIEW

In this chapter, I summarize the relevant literature related to the research contained in subsequent chapters. I review research covering the attention hypothesis and stock selection, mutual fund investor trading, the flow-performance relationship for mutual funds, and the family organizational structure in the mutual fund industry.

2.1 OVERVIEW OF RELATED ATTENTION LITERATURE

In this section, I review the finance literature related to the effect that attention has on the trading decisions of individual investors. While this topic is relatively young, there still exist several papers that relate closely to my research in Chapter 3.

One of the first studies to explore the market implications of investors' collective awareness of a particular stock is Merton (1987). Merton proposes a two-period model of capital market equilibrium in which every investor knows only about a subset of the available securities. A key assumption of his model is that an investor will only include a security in his or her portfolio if he or she knows about that security. Merton shows that, under this assumption, expected returns will tend to be lower on better-known firms with relatively larger investor bases.

While Merton's seminal model assumes that informed investors trade only in stocks in which they have relevant information, later studies relax that assumption and allow that investors may trade in stocks simply because those stocks capture their attention. One of the earliest papers to recognize the importance of attention in the trading of individuals is Lee (1992). In a study of firms listed on the NYSE in 1988, Lee separates trades based on trade size and examines the directional volume of large and small trades in response to earnings surprises. In a rather puzzling result, the author finds that small trades, which serve as a crude proxy for individuals' trading, exhibit unusually high buying regardless of the direction of the earnings surprise. In discussing potential explanations for this result, Lee recognizes the possibility that "it isn't the news content of the event per se, but the attention given to the firm, that triggers the small trade response we observe."

Lamont and Frazzini (2007) also examine the effect of attention related to earnings announcements. The authors document that stock prices rise around scheduled earnings announcements and that this premium is strongly related to trading volume surges. Lamont and Frazzini further show that stocks experiencing the largest premiums also experience the highest levels of small investor buying, suggesting that the premium may be caused by the buying of small investors whose attention is captured by the announcement.

Similarly to Lee, Odean (1999) also considers the effects of attention in discussing the individual trading patterns documented in his research. The primary focus of Odean's paper is whether individual investors at a large discount brokerage trade too much, resulting in poor portfolio performance after accounting for transactions costs.[1] Among the interesting trading patterns that Odean discovers is that individuals buy stocks that have experienced greater absolute price changes in the past. In his discussion of the results, the author carefully considers possible explanations for this behavior. Chief

[1] Note that Odean (1999), as well as Barber and Odean (2006) and Graham and Kumar (2006), explore the same dataset of individual trades at a large discount brokerage that I describe in Chapter 3.

among these explanations is the idea that investors are limited in their ability to evaluate the over 10,000 stocks available in the market and instead only consider purchasing stocks to which their attention has been drawn. Odean also proposes that attention will affect individuals' buy and sell decisions asymmetrically, as individuals are reluctant to sell short and need not rely on any simplification to help them choose among the stocks already held in their portfolios when they need to sell. Thus, in his discussion, Odean formulates the basis of what is now known as the attention hypothesis. The author does not, however, perform any additional tests to confirm his conjectures about the role of attention in the purchase decisions of individuals.

Following Odean (1999), several papers have conducted studies in various contexts which confirm that individuals are net buyers of attention-grabbing stocks. In a study of the existence of dividend clienteles among individual traders, Graham and Kumar (2006) present evidence confirming the importance of attention for individual stock choice. Specifically, the authors show that individual investors as a group exhibit net buying on the day of and days following a dividend announcement, even though there is no consumption or tax-motivated reason for doing so. Moreover, the authors find that this attention effect is most pronounced for older and low-income investors, the two investor groups that display the strongest preferences for dividends in the study.

Two studies motivated as tests of the attention hypothesis and, hence, most closely related to my own are Barber and Odean (2006) and Seasholes and Wu (2007). Barber and Odean (2006) test whether individual traders at a large discount brokerage are net buyers of attention-grabbing stocks. Specifically, the authors document that individuals are net buyers of stocks that receive news mentions, exhibit abnormally high trade volume or experience large one-day price movements on the previous day. The research in Chapter 3 of this dissertation adds to this line of research in two distinct ways. First, unlike Barber and Odean, I examine attention-grabbing events that need not be associated with pertinent information or extreme price or volume fluctuations. By demonstrating that individuals respond similarly to attention events that are not closely tied to underlying information events, I confirm that the behavior documented by the authors is in response to attention, per se, and not a systematic response to underlying information that causes the attention-grabbing events. Second, I expand beyond the scope of the Barber and Odean study to show that attention impacts not only individual trading behavior, but also stock prices, leading to predictable short-term returns following specific attention-generating events.

A closely related paper in which the authors consider the implications of attention-motivated trading on stock price efficiency is Seasholes and Wu (2007). In their paper, the authors study instances in which stocks on the Shanghai Stock Exchange reach their upper daily price limits.[2] The authors note that these events are characterized by high returns, high volume, and news mentions. Consistent with the findings of Barber and Odean (2006), the authors find that active individual investors in this market are net buyers of stocks that reached their upper price limits on the previous day. Furthermore, the authors demonstrate that stock prices exhibit initial price increases following these events, which are followed by subsequent reversals over the following week. However,

[2] The Shanghai Stock Exchange imposes daily limits on the price movements of listed stocks.

as the authors recognize, their return findings are plagued by a censoring bias, created by the fact that the exchange allows trading to continue once a stock hits its upper price limit, but transactions prices may not exceed the limit. My research is differentiated from Seasholes and Wu in four ways. First, our samples differ substantially; Seasholes and Wu study the Shanghai Stock Exhange, while I focus on Nasdaq. Second, I am able to test trading and return patterns around both 52-week highs and lows, while data availability constrains Seasholes and Wu to studying only events involving upper daily price limits. Third, unlike the results in Seasholes and Wu, my return results do not suffer from any institutional restrictions that bias the results towards finding positive abnormal returns on days following attention-grabbing events. Finally, upper daily price limit events necessarily require that stocks have experienced large positive one-day returns on the event day; thus, the authors cannot disentangle the effects of attention generated by upper price limit events from the returns themselves or from the underlying information that is likely to trigger such returns. By studying 52-week highs and lows, which do not necessarily coincide with large returns, I am able to disentangle the two effects.

Finally, my work is also related to Huberman and Regev (2001), who examine the case of a non-informative event which had a dramatic effect on stock market prices. The authors document that stock prices for EntreMed soared in response to a *New York Times* article on the potential development of new cancer-curing drugs, even though the possible breakthrough had already been reported in multiple journals and newspapers more than five months earlier. While the authors' research centers around a single nonevent, I show that enthusiasm related to attention-grabbing, uninformative events that occur frequently, such as 52-week highs and lows, can still affect stock prices.

2.2 OVERVIEW OF RELATED MUTUAL FUND LITERATURE

In Chapter 4, I examine the effect of mutual fund family reputation on the fund choices of individuals at a large discount brokerage house. Although this study is the first to empirically investigate the importance of reputation for the mutual fund selections of investors, there is survey evidence in the literature suggesting that reputation may be an important factor. Capon, Fitzsimons and Prince (1996) analyze self-reported survey data from 3,386 mutual fund investors and find that investors consider attributes other than risk and return when making mutual fund investment decisions. Interestingly, two of the three most important selection criteria for surveyed investors were fund manager reputation and the number of funds offered by the family.[3] Similarly, a study by the Investment Company Institute (1997) finds that fund family reputation was the third most important piece of information considered for the most recent fund purchase, after risk level and total return. In a survey of financial advisors, Jones, Lesseig and Smythe (2005) report that fund manager reputation was the fifth most important determinant of fund choice.

In addition to this survey evidence, several other papers relate closely to my research. For the remainder of this section, I review other works from the mutual fund literature that are most related to the research in Chapter 4. Specifically, I summarize relevant papers that examine the trading behavior of mutual fund investors, the flow-

[3] The most important selection criterion was the fund's investment track record.

performance relationship for mutual funds, and the organizational structure and incentives of mutual funds and mutual fund families.

2.2.1 Mutual Fund Investor Trading

A number of recent works have studied the behavior of investors in the context of mutual fund selection. In this subsection, I summarize the studies that are most relevant to the mutual fund research conducted in Chapter 4.

To begin, several papers are related to my work because they focus on the factors that affect the observed mutual fund trades of individual investors. Two such papers are Barber, Odean and Zheng (2005) and Ivkovic and Weisbenner (2006), both of which exploit the same database of individual investors at a large discount brokerage that is examined here. Barber, Odean and Zheng find that mutual fund flows respond negatively to front end load fees, but have no discernable relationship with operating expenses. The authors submit that investors want to avoid fees, but pay more attention to salient, one-time fees like front loads. Ivkovic and Weisbenner find that, in contrast to the disposition effect displayed in individuals' stock holdings, individuals are reluctant to sell mutual funds that are winners and are more willing to sell losers. The authors also find that individuals' fund inflows react to relative returns while fund outflows, particularly in taxable accounts, are sensitive to absolute returns. Finally, Johnson (2004) analyzes a proprietary database of individual shareholders in one no-load mutual fund family and explores investors' horizon predictability and the costs related to short-horizon versus long-horizon investors.

Another relevant strand of the mutual fund literature examines whether investors make sound investment decisions when selecting funds. At present, these works have found evidence that both support and refute the notion that investors skillfully select their fund holdings. For instance, Zheng (1999) examines the question of whether investors have ability when it comes to picking mutual funds. The author finds that funds that receive new money significantly outperform funds that lose money, and only a portion of this effect can be explained by chasing past winners. In aggregate, the author reports that an implementable portfolio strategy based on flow information for small funds can reliably earn excess returns. On the other hand, Elton, Gruber and Busse (2004) show that investors appear to make irrational decisions when investing in S&P 500 index funds. Even though the products are basically homogenous, the authors discover significant heterogeneity in fees and net performance among the index funds in their sample. The authors find that a naïve investment strategy based on predicted performance outperforms the actual index fund holdings of investors. In the results presented in Chapter 4 of this dissertation, it appears that individual investors do exhibit some skill when selecting mutual funds, as individuals' portfolios earn positive average objective-adjusted monthly returns.

Another related topic is the cost imposed by constraining the set of choices available to mutual fund investors. Elton, Gruber and Blake (2006) and Elton, Gruber and Green (2007) both quantify the costs associated with limiting investors' mutual fund options. In a study of the 401k offerings of over 400 plans, Elton, Gruber and Blake find that only 53% of plans offered an adequate set of mutual fund options. The authors further show that this limitation has a significant negative effect on the terminal wealth of constrained 401k participants over a 20-year investment period. Elton, Gruber and Green address the risk associated with confining mutual fund investments to a single fund family. The authors find that fund correlations are higher within versus across fund families. This correlation stems from a propensity for funds within a family to hold the same stocks and exhibit similar exposure to macro risk factors. The authors estimate that an investor who chooses to invest within the same fund family rather than without would need to be compensated with an additional 50 to 70 basis points in return to maintain the same Sharpe ratio. Although the authors give several reasons why investors may confine mutual fund investments to a single family, data limitations prevent them from documenting the extent to which these investment practices actually take place. There are reasons to believe, however, that this cost related to diversification loss for concentrated mutual fund investors may be mitigated if investors choose to concentrate in families for which they have an informational advantage. For example, using the same sample of individual investors studied here, Ivkovich, Sialm and Weisbenner (2006) find that households that concentrate their common stock investments in a few stocks outperform households with more diversified accounts. This outperformance is strongest for small and local stocks, supporting the hypothesis that concentrated individual investors successfully exploit information asymmetries.

2.2.2 Mutual Funds and the Flow-Performance Relationship

In Chapter 4, I examine the effect that proxies for fund family reputation have on the flow-performance relationship for mutual funds. A wealth of previous research has explored the relationship between mutual flows and past performance. Although an exhaustive list of every study conducted in this area is beyond the scope of this review, below I discuss several of the papers that most closely relate to my research.

A seminal paper relating to flow-performance sensitivity in mutual funds is Sirri and Tufano (1998). Sirri and Tufano conduct an extensive study of fund flows into and out of mutual funds and document the asymmetric nature of the relationship between flows and past performance—investors chase high past returns far more strongly than they flee from poor returns. Goetzmann and Peles (1997) similarly find that the positive relationship between past performance and flow is confined to the top quartile of past performers. Del Guercio and Tkac (2002) show that the relationship between flows and past returns differs for mutual funds and pension funds, as pension fund flows severely punish poor performance and do not strongly chase good performance. Theoretical justification for the documented flow-performance relationship is provided by Berk and Green (2004), who present a rational model that predicts flows that chase performance even though mutual fund managers' returns do not exhibit persistence or outperform passive benchmarks. Gallaher, Kaniel and Starks (2006) show that, in addition to the past performance of the individual fund, flows also exhibit a similar positive relationship with the past return of the entire fund family.

Of particular relevance to my research are studies of how fund characteristics affect flow-performance sensitivity. Sirri and Tufano find that the flow-performance relationship is stronger for high-fee funds, while large family membership and media attention have no significant effect on the sensitivity of flow to past performance. Huang, Wei, and Yan (2006) find that large family affiliation increases flow-performance sensitivity to medium performance, but reduces sensitivity to superior performance. In contrast, the results in Chapter 4 show that the flow-performance sensitivity of individual mutual fund investors is significantly weaker for funds belonging to large families. It may not be surprising, however, that the individual investors in my discount brokerage sample behave differently from the flows observed by aggregate investors. Indeed, both of the papers discussed above highlight the role that family membership plays in lowering participation costs for investors. In the dataset that I analyze, fund family membership does not substantially alter the structure of participation costs, since investors are buying the funds through omnibus accounts.

Huang, Wei and Yan (2007) examine the effect of past return volatility on the flow-performance relationship. The authors document that flow-performance sensitivity is significantly weaker for funds with more volatile return histories. The authors also find that this dampening effect is more pronounced for younger funds. This result is consistent with investors relying more heavily on past performance to form their expectations when their priors are more uncertain, as predicted by the investor learning hypothesis. This hypothesis also relates well to my findings that membership to larger and older families reduces flow-performance sensitivity. My results indicate that fund family reputation also serves as a mechanism by which uncertainty in investors' priors is reduced, thereby easing investors' reliance on past performance in the expectation formation process.

2.2.3 Mutual Fund Organizational Structure

Finally, my research is related to studies of the incentives organizational structure in the mutual fund industry. In this subsection, I review relevant studies related to organizational structure, focusing primarily on research that investigates the incentives of management companies and fund managers to organize into large families of funds.

Massa (2003) proposes that mutual fund families can lower effective fees by offering many differentiated funds among which investors can freely allocate their capital. The author shows that, as a result, funds belonging to large families attract investors with shorter horizons or who value frequent turnover. The author further proposes that because fund families can differentiate their products through channels other than performance, it may not always be optimal for funds to maximize performance in order to increase market share and limit competition. This intuition is supported empirically by the finding of a significant negative relationship between performance and the degree of product differentiation.

Nanda, Wang and Zheng (2004) provide an additional motivation for family organization by documenting positive spillover effects for funds in families that produce star funds. The authors show that funds with higher variation in investment strategies are more likely to produce star funds and hypothesize that funds with less ability may pursue performance-destroying strategies to produce stars. Khorana and Servaes (2005) also highlight the benefits of increasing the number of offerings within a single family. The authors examine the importance of competition and the determinants of market share in the mutual fund industry and find that funds can increase market share by altering fee structure or by differentiating their product. In particular, improving performance and introducing more funds relative to the competition can aid in increasing market share. My research builds on these papers by offering an additional rationale for organization into large fund complexes.

CHAPTER 3: THE EFFECT OF ATTENTION ON STOCK RETURNS AND INDIVIDUALS' STOCK CHOICES

3.1 INTRODUCTION

The attention hypothesis, as first postulated by Lee (1992) and further developed by Barber and Odean (2006), predicts that individual investors will be net buyers of stocks that "catch their eyes." Empirical evidence examining the trading behavior of individual investors has provided strong support for this hypothesis.[4] In this chapter, I examine the question of whether the trading behavior of these attention-driven investors influences aggregate security prices and induces stock return predictability following two particular attention-grabbing events: 52-week highs and lows.

Recent literature in the arena of behavioral finance has uncovered numerous examples of suboptimal investment behavior on the part of individual investors.[5] In many cases, this boundedly rational behavior is the result of systematic behavioral biases that can induce correlations in the trading patterns of individuals. In a market with limits to arbitrage, if the trades of these biased market participants are concentrated, then their behavior can exert price pressure on the securities in which they trade. Furthermore, if the direction and timing of these trades are predictable, then one might be able to forecast future returns. In order to test for the existence of this price impact, the researcher first needs a theory of investor behavior that predicts active trading, as well as the approximate timing and direction of those trades. The attention hypothesis, which is

[4] See Section 3.2 for a review of the empirical evidence supporting the attention hypothesis.
[5] See Hirshleifer (2001) for a thorough review of this literature.

described in detail in Section 3.2, provides such a theory, as it predicts that individual investors will be net buyers of stocks that have recently captured their attention.

In this paper, I contribute to an understanding of how attention affects individuals' investment decisions by examining whether this predicted net buying by attention-driven investors leads to discernable price pressure and hence positive and predictable short-run returns following attention-grabbing events. Unfortunately, events used previously in the literature to define the set of attention-grabbing stocks, such as extreme returns, high trading volume, news mentions, dividend announcements and earnings surprises, make testing and quantifying the effect of attention on future returns difficult. In particular, it is exceedingly challenging to establish causality in such a study because each of the attention-grabbing events listed above either explicitly or implicitly indicates the arrival of new information to the market. Thus, even if the researcher could document short-run excess returns following one or more of these events, it would be very difficult to prove that these returns are caused by price pressure resulting from increased attention rather than from some systematic, and perhaps inefficient, incorporation of the revealed information into the price.[6]

To address this difficult causality issue, I introduce the novel idea of defining the set of attention-grabbing stocks as securities that hit new 52-week highs or 52-week lows. This new classification offers several key advantages in the tests that follow. First, unlike other events that have been studied in the attention hypothesis literature, 52-week highs

[6] For example, Barberis, Shleifer and Vishny (1998) predict conservatism, or under-reaction, in response to informative signals, due to an investor's overconfidence in his prior assessment of firm value.

and lows do not necessarily convey any new information about the fundamentals of the underlying firm; they are simply a measure of the current intraday price relative to an arbitrary historical benchmark. In spite of this fact, however, 52-week highs and lows continue to garner considerable attention. These events are widely followed and are frequently and saliently reported by financial media. Since 52-week highs and lows receive substantial attention without necessarily being informative, these events have considerable value when trying to disentangle the effects of attention and information on future returns.

Second, by examining the effect of 52-week highs and lows on returns, I am able to explicitly control for the effects of past return, trading volume and volatility, variables which might be linked to future returns for reasons other than attention.

Lastly, this new classification of attention-grabbing stocks allows for the experiment of comparing stock returns following 52-week highs and lows to returns following similarly-defined n-week highs and lows that are not simultaneously 52-week highs or lows.[7] As n approaches 52, the properties of n-week high and low stocks should be almost identical to 52-week high and low stocks, with the lone difference being the considerable attention paid to 52-week events. The ability to form two samples that differ only in the amount of attention paid to them provides a unique glimpse into the possible causality of any abnormal returns that are discovered. Further discussion regarding the merits of examining 52-week highs and lows is found in Section 3.3.

[7] For instance, I compare abnormal returns following 52-week highs to those following 49-week highs.

The primary purpose of this paper is to test whether the predicted net buying from individual investors following attention-grabbing events affects security prices in a predicable and measurable way. Before those tests are conducted, however, it is first necessary to establish that investors are, in fact, strong buyers following the attention-grabbing events studied in this paper, i.e. 52-week highs and lows. Thus, the first step in my empirical design is to document the trading patterns of individuals in stocks that reach 52-week highs and lows. Using individual household trading data from a major discount brokerage and following the methodology described in Barber and Odean (2006), I find that the buy-sell imbalances for stocks that reach new 52-week highs or lows on the previous day are significantly more positive than for all other stocks and days in the sample. This finding obtains regardless of whether the buy-sell imbalance measure is defined using the number of shares traded or the value of the trades. In sum, the individual trading study reveals that 52-week high and low events do draw attention from sample individuals, resulting in significantly more positive net purchases by individual investors in these stocks.

Having established that individual investors are buyers, on average, following 52-week highs and lows, I next turn to the question of whether this attention-driven buying exerts measurable upward pressure on prices. In daily Fama-MacBeth regressions where the dependent variable is the market-adjusted return on stock j and the independent variables are dummy variables that equal one if stock j hit a 52-week high or low on the prior day, I find that stocks reaching new 52-week highs (lows) experience significant abnormal next-day returns of 30.7 (9.4) basis points. After controlling for autocorrelation in returns, I estimate the effect of a 52-week high on subsequent daily returns to be 22.1 basis points, while the effect of a 52-week low is estimated to be 29.9 basis points.[8] This key result, that 52-week highs and lows both have positive and statistically significant effects on next-day stock returns, is robust to the inclusion of additional explanatory variables into the regression to control for event-day, stock-level turnover, abnormal turnover, liquidity and volatility.

After documenting that 52-week high and low events do have a significant, positive effect on next-day returns, I next turn to the issue of causality. A powerful result emerges from comparing returns subsequent to 52-week highs and lows to returns following similarly-defined highs and lows over slightly shorter reference periods. In this experiment, I examine the effect that 47-, 49-, or 51-week highs and lows, which are not simultaneously 52-week highs or lows, have on next-day returns. Interestingly, the tests reveal that these alternatively-defined high and low events have no significant impact on next-day returns. This result is particularly revealing, since the properties of the 47-, 49-, or 51-week high and low stocks should be almost identical to the properties of the 52-week high and low stocks; the only difference is that 52-week highs and lows are prominently followed and highly publicized, while highs and lows defined over different reference periods are not. Thus, these results strongly support the hypothesis that the significant, positive effects of 52-week highs and lows on future returns are the direct result of the extra attention paid to these particular events.

[8] Both coefficient estimates are highly statistically significant, with corresponding t-statistics of 6.18 and 7.34, respectively.

Next, I attempt to show that it is the mere incidence of a 52-week high or low, rather than some underlying new information, that causes the 52-week high and low effects. To that end, I exclude from the sample any 52-week high or low that occurs within plus or minus three days of the underlying firm's earnings announcement. In addition, I rank all 52-week highs (lows) based on event-day absolute return, turnover, and intraday volatility and exclude any high (low) event that ranks in the top half of any of the three categories. Repeating the predictive regressions for this subsample which includes only those 52-week highs and lows that occur on relatively uneventful days, I still find that 52-week highs and lows have positive and significant impacts of 21.5 and 17.9 basis points, respectively, on next-day returns.

Additionally, I test whether the 52-week high (low) effect is stronger for days when relatively fewer 52-week highs (lows) occur. If the 52-week high and low effects do indeed result from attention-generated price pressure, then the effects should be strongest when these two attention-grabbing events are most effective in narrowing investors' choice sets. Thus, I first separate all sample days into 10 groups based on the number of 52-week highs that occur on that particular day. I then estimate the predictive return regressions described above within each decile. Consistent with my conjecture, on days with the fewest occurrences of 52-week highs, I find that the effect of a 52-week high on a stock's next-day return is 25.1 basis points, as compared to a small and statistically insignificant effect of 4.58 basis points on days which experience the greatest number of 52-week highs. Sorting by the number of 52-week lows that occur on a particular day reveals a similar result. Within the decile corresponding to the fewest occurrences of 52-week lows, the effect of a 52-week low on next-day returns is a highly significant 34.8 basis points. This effect is significantly larger than within the decile with the highest number of 52-week lows, where the effect of a 52-week low on next-day returns is estimated to be -0.75 basis points.

Finally, I investigate the potential causal relationship between individual trading and the abnormal returns following 52-week highs and lows. Although shortcomings of the individual trade dataset limit statistical power, I provide evidence that the positive abnormal returns that follow 52-week highs are confined completely to the set of events that result in increased trading by individuals on the following day. For both 52-week highs and lows, the coefficient estimates are positive and economically large for events with positive individual buy-sell imbalances on the next day, while the same coefficient estimates are negative for events with negative next-day buy-sell imbalances.

The remainder of the chapter is organized as follows. Section 3.2 provides a more thorough explanation of the theoretical and empirical literature regarding the attention hypothesis. Section 3.3 further motivates using 52-week high and low events to define the set of attention-grabbing stocks. Section 3.4 describes the data, sample selection procedure, and variable definitions. Section 3.5 examines individual household trading behavior following 52-week highs and lows. Section 3.6 describes the empirical findings of return predictability following 52-week highs and lows. Section 3.7 conducts additional tests to ascertain the possible causality of the abnormal returns. Section 3.8 conducts robustness checks, and Section 3.9 discusses the economic significance of the findings. Section 3.10 concludes.

3.2 THE ATTENTION HYPOTHESIS: EVIDENCE AND INTERPRETATION

When considering which stocks to include in their portfolios, individual investors face the daunting task of selecting from the thousands of equities that are available in the market. Due to cognitive constraints and time limitations, individuals may rely on heuristics to narrow the pool of stocks under their consideration. Odean (1999) proposes that investors may simplify the search problem by simply choosing from among the subset of equities that have recently attracted their attention. Therefore, it is predicted that individual buy demand will be relatively higher for attention-grabbing stocks.

On the other hand, because individual investors are typically reluctant to sell short, they do not face the same search problem when it comes to selecting securities to sell.[9] Therefore, individuals are not forced to rely on simple heuristics when selling. Instead, these investors will only consider selling the stocks which they currently hold in their portfolios, which is typically a very small subset of the set of all available stocks.[10] Because the predicted increase in individual buy demand following attention-grabbing events is not offset by a similar increase in predicted individual selling, the attention hypothesis proposes that the net demand for stocks that catch the eye of individual investors will be positive.

Several empirical papers find support for the attention hypothesis. Lee (1992) finds that investors who place market orders of less than $10,000 are net buyers

---

[9] Barber and Odean (2005) report that only 0.29 percent of positions in their large discount brokerage dataset are short positions.
[10] Barber and Odean (2005) report that the mean household in their dataset held only 4.3 stocks in its portfolio; the median household stock holding was 2.61.

subsequent to both positive and negative earnings surprises, and this finding is later confirmed by Hirshleifer et al. (2003) using individual trade data.[11] In addition, Barber and Odean (2006) document that individuals are net buyers on days that a stock experiences high turnover, on days following extreme price movements, and on days after the underlying company has been prominently mentioned in the news. Graham and Kumar (2006) report that individuals tend to be net buyers of stocks following dividend announcements, a result which the authors attribute to the attention effect. Finally, in contemporaneous work, Seasholes and Wu (2007) find that individual investors on the Shanghai Stock Exchange initiate new positions in, and are net buyers of, stocks that reach their upper price limits on the previous day.[12]

3.3 MOTIVATION FOR STUDYING 52-WEEK HIGHS AND LOWS

As noted in the introduction, this study defines the set of attention-grabbing stocks as stocks that have recently hit a new 52-week high or 52-week low. In this section, I elaborate on the advantages of choosing these particular events when testing for the effect of attention on future stock returns.

First, 52-week highs and lows do not contain information about the fundamentals of the underlying company, per se. Certain events that have been used in previous studies to classify attention-grabbing stocks, such as news mentions or earnings announcements, are synonymous with the arrival of news to the market. Thus, while evidence of abnormal positive returns subsequent to these eye-catching information events would be consistent

[11] Lee classifies transactions in his data according to trade size under the assumption that trades under $10,000 are more likely to represent individual investors.
[12] The Shanghai Stock Exchange imposes daily limits on the price movements of listed stocks.

with attention-based buying impacting prices, it would be difficult to disentangle the attention explanation from models of price predictability based on the life cycle of news events. Fifty-two week highs and lows, on the other hand, need not be associated with any relevant information about the firm. For example, a stock that experiences a small price change on a day where no information is released could still reach a new 52-week high or low, particularly if the preceding day's closing price was already near to the previous high or low. In other words, even if stock price processes were driven entirely by random walks, one would still frequently observe stocks reaching new 52-week highs or lows. Admittedly, stock prices are affected by information, and 52-week highs and lows are likely to be correlated with informative events. In particular, stocks may tend to hit 52-week highs or lows on days characterized by extreme price movements or high volatility resulting from the arrival of news to the market. However, by defining the attention-grabbing set of stocks based on 52-week high and low occurrences, I am able to control for various firm-specific factors, such as extreme past returns and high turnover, that may be associated with information arrival, thus isolating the impact that attention, in the absence of new information, has on returns.

Second, although these events have no particular economic relevance, it is important to note that 52-week highs and lows are, in fact, attention-grabbing events. As a salient and highly publicized statistic, the incidence of a 52-week high or low is likely to catch the eyes of investors searching for stocks to purchase. Several leading financial publications such as *The Wall Street Journal*, *Investors Business Daily*, *Financial Times*, and the *South China Morning Post*, to name a few, print daily lists of the stocks that closed the previous day at a new 52-week high or low. Moreover, in *The Wall Street Journal*, among others, a bold arrow pointing up or down accompanies the listings of stocks that hit a new high or low the previous day, causing these securities to stand out from the other issues listed on the page. Furthermore, George and Hwang (2004) present evidence that investors use the 52-week high as a reference point against which they evaluate the potential impact of news.

Third, 52-week high and low events are prevalent in the data. During the 20-year sample period, there are approximately 343,000 52-week high events and 282,000 52-week lows. The large number of events found in the data allows for the necessary statistical power to detect and precisely estimate the impact of attention on daily stock returns.

Fourth, the impact of 52-week highs and lows can be precisely measured in historical data. Other attention-garnering events, such as earnings announcements, mergers and acquisitions, or company name changes, do not allow for a perfectly clean event study, since information regarding an upcoming announcement could be leaked to the market by insiders, impacting the price of the security prior to the actual event date. Fifty-two week high and low events are free from this problem, however, since no one knows ex-ante when one of these events will occur.

Finally, a firm cannot endogenously choose to reach a new 52-week high or low. Financial theory, e.g. Miller and Rock (1985), has demonstrated that endogenously chosen events such as dividend initiations can be used by the firm to signal unobservable quality measures to the market. Since firms cannot self-select their securities to hit 52-week highs or lows, these events reveal no additional information in a signaling model.

3.4 DATA DESCRIPTION

3.4.1 Data Sources

The data in this chapter come from three main sources. First, stock-level data is collected from CRSP and I/B/E/S. From the CRSP daily stock files, I collect the following information: closing price, closing bid and ask, intraday high and low prices, number of shares outstanding, and share volume. From I/B/E/S, I collect earnings announcement dates for all sample stocks.

The third data source contains detailed information on the trades of over 78,000 individual households at a large discount brokerage from January 1991 to November 1996.[13] For each trade, I collect the trade date, price, number of shares, eight-digit security CUSIP and whether the position is long or short. Trades are then aggregated over all households for each stock on each date.

3.4.2 Sample Selection

The twenty-year sample period runs from January 1986 through December 2005. To be included in the sample, a security must have data on closing bid, closing ask, and intraday high and low prices for at least 52 consecutive weeks. Also, stocks whose prices are below one dollar on the first trading day of any given year are excluded from the

[13] This data set is described in detail in Barber and Odean (2000, 2001, and 2002).

sample in that year, so as to eliminate the effect of the returns of penny stocks on the analysis.

Because this study analyzes stock price behavior around 52-week high (low) events, the securities under examination will almost surely have experienced only positive (negative) price movements on the event day.[14] Therefore, it is imperative to make a correction for bid-ask bounce, the negative serial correlation exhibited by securities' closing prices when measured at short-term intervals due to market microstructure effects. To eliminate the bid-ask bounce problem, I calculate daily returns using the midpoint of the closing bid and ask prices on each stock.[15] Since CRSP only reports data regarding closing bid and ask prices for Nasdaq stocks, the sample is comprised only of stocks listed on Nasdaq during the sample period. The resulting sample contains 10,627 firms and a total of 16.6 million daily observations.

3.4.3 Variable Definitions

Several variables used in the empirical analysis must be derived from the CRSP data. Daily returns are calculated as daily percentage change in the bid-ask midpoint. Size is calculated as closing price multiplied by shares outstanding. Turnover is defined as daily volume divided by shares outstanding. The bid-ask spread is the difference in the closing bid and ask prices, and the relative spread is defined as the bid-ask spread scaled

---

[14] While it is possible, for example, for a stock's intraday price to rise to a new 52-week high level and still close below the previous day's closing price, resulting in a negative return on a 52-week high event day, these instances are exceptionally rare.

[15] See Lease, Masulis, and Page (1991) for a more thorough description of the merits of using the bid-ask midpoint to correct for bid-ask bounce.

by the bid-ask midpoint. Daily volatility is defined as the difference in the intraday high and low prices divided by the closing price.

I also must identify 52-week high and 52-week low events in the data. A stock is said to have reached a 52-week high on any day when its highest intraday price exceeds the highest intraday price for the past 52 weeks for that stock. Similarly, a 52-week low occurs when a stock's intraday low price is smaller than the lowest intraday price experienced in the past 52 weeks. In later sections, I also make use of n-week high and low events ($n<52$). For reasons that are apparent in later analysis, a stock is defined to have reached an n-week high (low) on any day when its highest (lowest) intraday price exceeds the highest (lowest) intraday price for that stock over the past n weeks but does not exceed the highest (lowest) intraday price for the past 52 weeks.

Finally, I define buy-sell imbalance (BSI) and value buy-sell imbalance from the database of individual trades at a large discount brokerage house. To begin, I aggregate all households' trades for each stock and each day. Stocks are then sorted into groups based on common characteristics. The BSI measures the amount by which the number of buys of a particular group outpaces the number of sells, and it is expressed as a percentage of all trades made by individuals in those stocks. The value buy-sell imbalance is similarly defined, but measures the market value of buys and sells rather than simply the number of trades.

3.4.4 Summary Statistics

Table 3.1 provides year-by-year summary statistics of the sample. The table reports the number of sample stocks each year, as well as the average and median market capitalization, price and bid-ask spread. The table shows that the number of Nasdaq firms grew steadily during the first half of the sample, peaking in 1998 and steadily declining thereafter. Table 3.1 also shows that the average market cap grew steadily from 1986 until 2000, with a large jump coinciding with the peak of the well-documented tech bubble. Finally, the table demonstrates that bid-ask spreads have diminished during the later years of the sample, dropping 80% from 0.6 in 1993 to 0.12 in 2005.

Table 3.2 shows the number of observed 52-week highs and lows each year, as well as the average event-day return associated with each. Not surprisingly, Table 3.2 reveals that average returns are positive on days that stocks reach new 52-week highs and negative on days that are 52-week lows. The table also reports the annual return on the value-weighted Nasdaq index for each year in the sample. As one might expect, 52-week highs occur more frequently relative to 52-week lows in years in which the stock market performed well, while 52-week lows are relatively more common in years that the market performed poorly.[16]

3.5 INDIVIDUAL TRADING AROUND 52-WEEK HIGHS AND LOWS

As discussed in Section 3.2, the attention hypothesis predicts that individual traders will be net buyers of stocks that catch their collective eye. This section examines the trading behavior of individual households to determine whether individuals are net buyers or sellers following 52-week highs and lows.

[16] Indeed, I calculate the correlation between the annual difference in the number of highs minus lows and the annual return on the Nasdaq index to be 0.53.

The data examined in this section contains information from a large discount brokerage firm on the individual trades of 78,000 households from January 1991 through November 1996. Within this sample, I aggregate all households' trades for each stock on each day to calculate the total number of shares of each stock bought or sold by sample households on any given day. The resulting data set contains approximately 1.46 million aggregate daily stock trading observations. The intersection of this data set with the original sample set of firms described in Section 3.4 contains 407,054 observations, of which 31,969 correspond to 52-week high or low events.

If individuals are engaging in attention-driven buying once they become aware that a stock has reached a 52-week high or low, then we should observe individual net buying in these stocks on the day after the event.[17] To test for this pattern, I begin by dividing the individual trading sample into the three groups: observations for which the previous day was a 52-week high, observations for which the previous day was a 52-week low, and all other observations. Following the methodology in Barber and Odean (2006), I define the daily buy-sell imbalance for each group as follows:

$$BSI_{pt} = \frac{\sum_{i=1}^{n_{pt}} NB_{it} - \sum_{i=1}^{n_{pt}} NS_{it}}{\sum_{i=1}^{n_{pt}} NB_{it} + \sum_{i=1}^{n_{pt}} NS_{it}} \qquad (1)$$

[17] If some individuals become aware that a stock has reached a 52-week high or low before market close on the day of the event, then they may choose to trade the stock on the actual event day. Unfortunately, there is no way to tell what time of day the 52-week high or low occurred, because the intraday high and low prices from CRSP do not contain time stamps. Thus, the buying imbalance measure presented here only considers activity on the following day and may understate the actual size of the imbalance that results from attention-based buying.

where n_{pt} is the number of stocks in group p on day t, NB_{it} is the number of purchases of stock i on day t, and NS_{it} is the number of sales of stock i on day t. I also calculate buy-sell imbalances based on the market value of trades by substituting the value of stock i bought (or sold) on day t for NB_{it} (or NS_{it}) in the equation above.

Table 3.3 compares the average daily trade imbalances for each of the three groups described above. Panel A computes buy-sell imbalances based on the number of transactions, while Panel B computes imbalances based on the market value of each trade. Consistent with previous research using this data, Table 3.3 reveals that individual households in the sample are net buyers of stocks, on average, as mean net trading is positive for all three groups. More importantly, Table 3.3 shows that individuals' buy-sell imbalances are significantly more positive on days following either 52-week highs or 52-week lows. Panel A reveals that aggregate sample households have an average daily buy-sell imbalance of 11.54% (37.54%) for stocks that experience a 52-week high (low) on the previous day. For all other observations in the sample, the mean buy-sell imbalance is 5.42%. T-tests reveal that the difference in means of 6.12% (32.12%) between days following 52-week highs (lows) and all other days is significantly greater than zero at the 1% confidence level.

Similar results obtain when trade imbalance is calculated using the market value of trades rather than the number of trades. Panel B demonstrates that the average aggregate buy-sell value imbalance on days following 52-week highs (lows) is 7.8% (38.02%) higher than the average value imbalance for all other observations, a difference which is significant at the 1% level. In sum, the results in this section reveal that, consistent with the predictions of the attention hypothesis, aggregate individual households have significantly positive net purchases in stocks that reach 52-week highs or lows on the prior day. Moreover, these next-day positive trade imbalances in 52-week high and low stocks are significantly more positive than for all other stocks in the sample.

3.6 FIFTY-TWO WEEK HIGH AND LOW RETURN EFFECTS

The preceding section confirmed that sample individual traders significantly increase their net purchases of attention-grabbing stocks, i.e. stocks that reach new 52-week highs or lows on the previous day. The next step in my analysis is to ascertain whether this increase in attention results in positive pressure on the prices of these stocks, resulting in predictably positive returns immediately following both 52-week highs and 52-week lows.

3.6.1 Methodology

Recall that having the ability to test the effect of attention on future returns while controlling for various firm-level factors such as past return or turnover was one of the motivations for studying 52-week high/low stocks. In this section, I use multivariate regression analysis to isolate the effect of 52-week high or low events on future returns.

Specifically, I follow Fama and MacBeth's (1973) two-stage procedure and run daily cross-sectional regressions of the form:

$$\textit{Market-adjusted ret}_{j,t} = \alpha + \beta_1 \cdot high_{j,t-1} + \beta_2 \cdot low_{j,t-1} + \gamma \cdot X_{j,t-1} + \varepsilon, \qquad (2)$$

where *market-adjusted* $ret_{j,t}$ is the return on stock j on day t minus the daily return on the value-weighted Nasdaq index, $high_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1, $low_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week low at time t-1, and $X_{j,t-1}$ is a vector of t-1 measurable, firm-level control variables.[18] Statistical inference is performed using the time-series of the daily regression coefficients. All t-statistics are calculated using a Newey-West correction for serial dependence up to five lags.

3.6.2 Baseline Results

The Fama-MacBeth regression results are presented in Table 3.4. The first column of Table 3.4 presents the coefficient estimates when only the 52-week high and low dummy variables are included as independent variables. Consistent with the main hypothesis of this paper, I find that the incidence of a 52-week high (low) has a positive and statistically significant impact of 30.7 (9.4) basis points on a stock's next-day market-adjusted stock return.[19]

In the next regression specification, I include the return of stock j on day t-1, the event day, as an additional independent variable. The reasons for including past return in the regression are straightforward. The average daily market-adjusted return on the day that a stock reaches a 52-week high (low) is approximately 4.40% (-5.89%). If the return data exhibit strong positive (negative) one-lag autocorrelation, then the increase of 30.7 (9.4) basis points in excess returns following 52-week highs (lows) may arise

[18] Note that because I run daily Fama-MacBeth regressions, the market adjustment to the dependent variable will only affect the constant α in the regression. In unreported results, I find that using a market-model adjustment to the dependent variable has virtually no impact on the size or significance of the regression coefficients.
[19] Except where otherwise noted, statistical significance is measured at the 5% confidence level.

mechanically due to serial correlation, rather than resulting from the incidence of the 52-week high (low) event itself.[20]

The results in Column II of Table 3.4 confirm the importance of including past return as an explanatory variable in the daily return prediction regressions. The large and statistically significant coefficient on the past return variable indicates that sample returns display strong positive autocorrelation. More importantly, Column II reveals that both the 52-week high and low dummy variables remain positive and significant after accounting for this autocorrelation; the coefficient on the 52-week high variable is estimated to be 22.1 basis points, while the 52-week low coefficient is estimated to be 29.9 basis points.

Note that the estimate of the 52-week low variable in Column II more than triples from Column I. The reason for this increase is fairly intuitive; in a sample in which returns are positively autocorrelated, large negative returns on one day are likely to be followed by negative returns again on the next day. Thus, when the regression does not explicitly control for autocorrelation, the positive attention effect that a 52-week low exerts is partially offset by the negative effect that the serial correlation has on next-day returns. Once I control for autocorrelation in the regression specification, however, I am able to isolate the sizeable positive impacts that both 52-week highs *and* 52-week lows impart on future stock returns.

It is important to note that although the sizes of the coefficients on the dummy variables in Column II are relatively small in relation to the coefficient on past return, the effect of a 52-week high or low on next-day returns is actually quite large when

[20] I recognize that past returns may also have a nonlinear relationship with future returns. In Column VI, I include squared past return as an additional control variable.

compared to the effect of past returns. For instance, a 52-week low for stock j on day t-1 has the same positive impact on next-day returns, 29.9 basis points, as a return of 10.28% on day t-1 for the same stock. In Section 3.9, I discuss the economic significance of these results in more detail.

3.6.3 Trading Volume Controls

In Columns III and IV of Table 3.4, turnover is also included as an independent variable in the regression. The purpose of including turnover in the regression is to isolate the attention effect, as measured by the 52-week high and low dummies, from any effects that may be due to trading volume, such as the high volume return premium of Gervais, Kaniel, and Mingelgrin (2001).[21] Column III reveals that both the 52-week high and low variables remain positive and statistically significant after controlling for event-day turnover. The coefficient estimate on the 52-week high dummy in Column III is 17.3 basis points, while the coefficient on the low dummy is 25.5 basis points.

Although the turnover variable in Column III is intended to capture trading volume effects, it is not a perfect control for the high volume return premium of GKM. First, GKM use a nonparametric approach by sorting stocks into portfolios based on size and volume and then comparing portfolio returns, whereas I employ a regression approach. Another important difference is that GKM define their volume portfolios based on abnormal volume, i.e. the size of the formation day's volume relative to past volume for that particular stock. To better control for GKM's high volume return premium, I

[21] Gervais, Kaniel and Mingelgrin, henceforth GKM, find that stocks experiencing unusually high (low) trading volume over one day tend to appreciate (depreciate) over the course of the following month.

replace the simple daily turnover variable with a new control variable named abnormal turnover. Abnormal turnover for stock j on day t is defined as:

$$Abnormal\ turnover_{j,t} = turnover_{j,t} - \left(\sum_{n=1}^{49} turnover_{j,t-n} \right) / 49. \qquad (3)$$

The 49-day reference period for the abnormal turnover calculation is chosen for consistency with the GKM methodology.[22] Column IV of Table 3.4 displays the regression results when abnormal turnover is included as a control. Consistent with GKM, the coefficient on the abnormal turnover variable is positive and significant; however, the coefficient estimates for the 52-week high and low variables change very little in the new regression specification, and both remain positive and highly statistically significant. Thus, Columns III and IV indicate that the 52-week high and low effects remain strong determinants of future returns even after controlling for the effects of high event-day trading volume in a stock.

In the regression specification of Column V, I control for the interaction of event-day stock return and trading volume to allow for the possibility that return autocorrelations are different for high volume stocks and low volume stocks. The model of Blume, Easley, and O'Hara (1994) along with empirical work by Campbell, Grossman, and Wang (1993) and Conrad, Hameed, and Niden (1994) underscore the importance of the interaction between past return and trading volume in predicting future

[22] It should be noted that I use abnormal turnover as my trading measure, whereas GKM form their portfolios based on abnormal volume. Since GKM first sort stocks into size groups, their volume measures within groups are reasonably comparable without being scaled. Since I do not include controls for size, it is necessary for me to scale trading volume by shares outstanding so as to make the measure comparable across all stocks in the cross-section.

short-horizon returns. To control for this interaction, I divide sample stocks daily into two halves based on abnormal volume, which I define analogously to the abnormal turnover variable described above, except with volume in the place of turnover.[23] The *high volume dummy* variable takes a value of one if a stock's abnormal volume measure ranks in the top half of the sample on a given day, and it is set equal to zero otherwise. Similarly, the *low volume dummy* takes a value of one if the firm ranks in the bottom half for a given day. Finally, I replace event-day return in the regression with two variables defined as the interaction of event-day return with the *high volume dummy* and then the *low volume dummy*. The results of this regression are presented in Column V of Table 3.4. Once again, the two attention variables, 52-week high and 52-week low, remain positive and significant in the new regression. The coefficient for the 52-week high (low) variable is estimated to be 15.3 (26.1) basis points with a corresponding t-statistic of 7.4 (7.21).

3.6.4 Additional Controls

Having thus far established that the 52-week high and low dummies do not merely proxy for event-day return or turnover, I next attempt to rule out two other possible explanations for the significant and positive coefficients on the 52-week high and low variables. Because 52-week high and low events often occur on days marked by extreme positive or negative stock returns, it is possible that the 52-week high and low

[23] The choice of whether to use volume or turnover to best capture this interaction is difficult. Blume, Easley, and O'Hara (1994) suggest a volume measure; Campbell, Grossman, and Wang (1993) use a turnover measure; and Conrad, Hameed, and Niden (1994) use the number of transactions. For robustness, I also define an interaction variable using turnover rather than volume, and the results are unaffected.

variables simply proxy for the volatility of the stock price on the event day. To control for this possibility, I include squared event-day return into the regression specification.[24]

Another possible explanation for the 52-week high and low effects is liquidity. The price impact of order flow is more pronounced for illiquid stocks, which could cause these stocks' intraday price movements to be more extreme, resulting in more 52-week high and low events when stocks are relatively less liquid. Thus, I also include in the regression the relative bid-ask spread, a common metric for illiquidity, to test whether the positive 52-week high and low effects proxy for a premium paid to illiquid stocks.

The results from including relative bid-ask spread and squared return as additional independent variables are found in Column VI of Table 3.4. While the inclusion of these additional variables does appear to have an important impact on the magnitude of the coefficient estimate for the 52-week low variable, which is reduced from 25.5 basis points in Column III to 13.8 basis points in Column VI, both of the attention variables remain positive and statistically significant.

In summary, the regression results presented in this section show that stocks experiencing 52-week highs and lows exhibit positive abnormal returns, on average, immediately following these events. Furthermore, the 52-week high and low effects on next-day returns remain positive and significant even after controlling for the effects of past return, high volume, volatility and liquidity. These results are consistent with the

[24] If extreme returns were indeed driving the 52-week high and low effects, it would not necessarily contradict the attention story, since large one-day price movements are certainly attention-grabbing events. Indeed, financial media, such as the *Wall Street Journal,* regularly publish tables detailing the day's largest percentage price gainers and losers. However, if 52-week high and low events have no effect on returns beyond what is explained by the volatility of the event-day return, then a study which focuses on these particular events has little value.

hypothesis that 52-week highs and lows cause a predictable increase in attention, resulting in positive price pressure and predictable short-term excess returns immediately following these events.

3.7 INVESTIGATING CAUSALITY

In this section, I perform several tests to determine whether the excess returns following 52-week highs and lows documented in Table 3.4 do indeed result from increased attention. First, I test whether the effects of 52-week highs and lows on future returns differ from the effects of similarly-defined highs and lows over slightly shorter reference periods. If the attention associated with 52-week highs and lows is driving the 52-week high and low effects, then a similar effect should not exist for highs and lows over different reference periods that are not as highly publicized. In the second test, I attempt to show that the 52-week high and low effects exist even for stocks for which nothing noteworthy happens on the event day other than the 52-week high or low itself. Next, I test whether the 52-week high (low) effect is stronger for days on which relatively fewer 52-week highs (lows) occur. Finally, I investigate the potential causal relationship between individual trading and the return predictability following 52-week highs and lows.

3.7.1 Comparison to Highs and Lows of Different Reference Periods

The results in Columns II through VI of Table 3.4 control for a number of event-day stock-level characteristics which one might expect to behave differently on the day that a stock reaches a 52-week high or low. For instance, Column III controls for return and turnover on the event day, because one might reasonably assume that stocks would have higher (lower) than average returns and higher than average turnover on the days that they hit 52-week highs (lows). Another characteristic that is common among stocks reaching new 52-week highs (lows) is a long-term run-up (run-down) in price, which is not captured by the event-day controls in Table 3.4. It is possible, then, that the effects captured by the two attention variables might exist for any stocks with long-term return histories similar to stocks that reach new 52-week highs or lows. If, in fact, it is the return history, rather than the event itself, which causes the 52-week high and low effects, then we should find similar effects for stocks that reach, say, 49-week highs or lows. Alternatively, if the effect is unique to stocks that reach 52-week highs or lows due to the attention associated with these events, then we should find no effect for 49-week high or low events.

To test this possibility, I define two new independent dummy variables in Column III of Table 3.5. The *49-week high*$_{j,t-1}$ *(low)* variable is a dummy that takes a value of 1 if stock j hits an 49-week high (low), but not a 52-week high (low), on day t-1. Stocks for which the 49-week high (low) variable has a value of one should be comparable to the stocks for which the 52-week high (low) dummy is one, except that the 49-week stocks do not receive the same publicity as their 52-week counterparts; however, Column II shows that the 49-week high or low stocks do not experience significant excess returns on the day following the event.[25] The impact of reaching a new 49-week high (low), which is not also a 52-week high (low), on next-day returns is a statistically insignificant -9.7

[25] Column I of Table V is identical to Column VI of Table 4. It is included in Table 5 simply for the purpose of comparison with Columns II and III.

(5.8) basis points. This striking result lends strong support to the attention explanation for the 52-week high and low return effects. Although there is nothing magical about the 52-week reference period, 52-week high and low data are collected and reported by the financial media, and new 52-week high and low events are given considerable attention. As a result, these events have a significant positive impact on stock returns as attention-driven investors become net buyers of these stocks. In contrast, events that are almost identical in every aspect except that they are not widely publicized, such as 49-week highs or lows, do not appear to have any effect on these stocks' future returns.

One additional comment regarding the results in Column III of Table 3.5 is in order. The 49-week reference period is chosen for comparison because it is close enough to the 52-week period to make the two periods easily comparable, while still leaving enough time between the two periods to allow for a large number of events in the comparison set. Note that since the test is only meaningful when the alternative reference-period comparison set contains events that are not also 52-week highs or lows, the closer the alternative reference period is to 52 weeks, the fewer events will be in the comparison set.[26] Admittedly, though, the choice of 49 weeks is quite arbitrary. For robustness, I perform similar regressions where the high and low dummies are defined using 47- and 51-week reference periods. The results are presented in Columns II and IV of Table 3.5. The findings are quantitatively and qualitatively similar to those for the 49-week high and low; the coefficient on the 47- and 51-week high dummies are negative

[26] Of course, this is also the reason that the alternative reference period must be shorter than 52 weeks. For instance, any stock which hits a new 55-week low has, by definition, also reached a new 52-week low.

and insignificant, while the coefficients on the corresponding low dummies are positive, but small relative to the 52-week low variable, and statistically insignificant.

3.7.2 Excluding Probable News Events

Recall from Section 3.3 that one of the motivations for defining attention-grabbing stocks using 52-week highs and lows is that these events do not contain information about the fundamentals of the underlying firm, per se. Thus, this classification offers several advantages when it comes to the difficult task of disentangling the effects of attention and information in the data. However, while these 52-week events do not necessarily coincide with the arrival of news to the market, it would be unreasonable to believe that they do not sometimes occur in response to new information. For instance, a positive earnings surprise could cause a stock to experience a large and immediate price increase, triggering a new 52-week high on the day of the announcement. While I have argued that 52-week highs are not necessarily informative, it is obvious that in this example the new 52-week high is driven by new information. In this section, I try to identify those 52-week highs and lows that are most likely to be driven by information. I then remove these events from the sample and test whether the abnormal returns following 52-week highs and lows persist for the remaining observations.

As a first step, I exclude from the sample any 52-week highs or lows which occur within plus or minus three trading days of an earnings announcement. Some of the most common and widely analyzed releases of information to the market are quarterly and annual earnings announcements, and numerous studies have shown financial markets to react strongly to the information contained in these releases. Out of a total of 624,862 occurrences of 52-week highs or lows in the sample, 58,136 occurred within plus or minus three trading days of an earnings announcement date.[27] These events are excluded from the sample in the results reported in Table 3.6.

Although I am able to identify and exclude events that occur near earnings announcements, there are scores of other information events that are more difficult, or even impossible, to reliably identify and exclude from the sample. Although it would be extremely difficult to search the newswires for each of the roughly 625,000 52-week events identified in the 20-year sample period, there are reasonable ways to separate the events according to the likelihood that an informative event caused the 52-week high or low. For example, consider two stocks, A and B, which hit 52-week lows on the same day. On the day of the event, Stock A, which has been gradually approaching the 52-week low in recent weeks, has a negative return of 1.5% on below-average trading volume. Stock B, on the other hand, experiences unusually high trading volume and large fluctuations in intraday price and closes at a price that is down 10% from the previous close. In this example, it seems much more likely that the 52-week low experienced by Stock B was driven by underlying information.

Applying this logic, I sort all 52-week high events into two equal-sized groups ranking on absolute event-day raw return. I also perform similar independent sorts ranking on event-day turnover and event-day volatility.[28] All 52-week high events that

[27] The mere fact that less than 10% of the 52-week highs and lows in the sample occur within plus or minus three days of a quarterly or annual earnings announcement hints that these events are not driven by information.
[28] I define daily volatility as the difference in the intraday high and low prices divided by the closing price.

rank in the top 50% for one or more of the three sorts are labeled as “probable news events” and are excluded from the sample. The same procedure is then performed for 52-week low events. The remaining sample contains 68,652 52-week low events and 94,057 52-week high events. If the excess returns following 52-week highs and lows are caused by attention, independent of true information, then the return effects should remain just as strong within this subsample of relatively “uneventful events.”

Table 3.6 reports Fama-MacBeth regression results after excluding 52-week highs and lows that occur within plus or minus three days of an earnings announcement or that are determined to be probable news events. Interestingly, the coefficient estimates for both the 52-week high and low dummies are actually larger within this subsample. The 52-week low dummy increases from a value of 13.8 basis points in Column VI of Table 3.4 to 17.9 basis points in Table 3.6, while the coefficient on the high dummy increases by approximately 50 percent from 14.6 basis points to 21.5 basis points.

The results in this section reveal that the abnormal returns following 52-week highs and lows exist, and in fact appear to be stronger, even for the subset of events that are least likely to be caused by the arrival of firm-specific information to the market. These results provide two important insights into what may cause the 52-week high and low return effects. First, these results show that the effect is not likely driven by the fact that 52-week highs and lows could be correlated with information arrival. Second, by showing that the effect obtains even for event-days on which the stock price and trading activity are relatively uneventful, these results provide additional evidence that it is the 52-week high or low event itself, rather than other irregular price or volume characteristics that are common to stocks that experience 52-week highs or lows, that is driving the abnormal returns.

3.7.3 Examining the Daily Number of Highs or Lows

The key assumption underlying the attention hypothesis is that individuals rely on attention-grabbing events to narrow their consideration set when they decide to invest in stock. One would expect that the resulting attention-driven buying, and any abnormal returns it may cause, would be strongest when the attention-grabbing event being considered does the best job of focusing investor attention on a select group of stocks. For instance, if 100 stocks hit new 52-week highs on a particular day, then individuals wanting to purchase stock may not find focusing on 52-week high stocks to be an effective means of narrowing their stock search. Aggregate individual investors may spread their purchases among these 100 stocks, or they may pursue other methods that more adequately limit their consideration sets. On the other hand, if only one stock experiences a 52-week high on a given day, then this event is extremely salient and highly effective in narrowing investors' choice sets. Thus, if the 52-week high (low) effect does result from attention, then this effect should be stronger for days with relatively fewer 52-week highs (lows).

To test this hypothesis, I first separate all sample days into 10 groups based on the number of 52-week highs that occurred on the previous day. Within each decile, I estimate the Fama-MacBeth regression described by equation (2), including each of the control variables used in Table 3.6. Panel A of Table 3.7 reports the Fama-MacBeth estimates for $\beta_{1,}$ the coefficient for the 52-week high dummy variable, for each decile.

Notice that the effect of the 52-week high variable on next-day returns is generally larger on days when fewer 52-week highs occur. Indeed, the 52-week high effect is an economically and statistically significant 25.1 basis points for the days with the fewest 52-week highs, as compared to an insignificant 4.58 basis points for the days with the most 52-week highs. This difference of 20.51 basis points between the bottom and top deciles is significant at the 10 percent confidence level.

Next, I repeat the same procedure for 52-week lows. I sort all days into 10 groups based on the number of 52-week low occurrences and then examine the effect of a 52-week low on the following day's returns within each decile. Panel B of Table 3.7 reports the Fama-MacBeth estimates for $\beta_{2,}$ the coefficient for the 52-week low dummy variable, for each of the deciles sorted on the daily number of 52-week lows. As with the 52-week high effect, the 52-week low effect is stronger when fewer 52-week lows occur on a particular day. The 52-week low effect is an economically and statistically significant 34.78 basis points for the days with the fewest 52-week lows, as compared to an insignificant -0.75 basis points for the days with the most 52-week lows. This difference of 35.54 basis points between the bottom and top deciles is significant at the five percent confidence level.

The results in this section are again highly consistent with the hypothesis that the abnormal returns following 52-week highs and lows are the result of increased attention paid to these stocks. On days when 52-week highs (lows) do the best job of focusing an investor's attention on a concentrated group of stocks, i.e. when a relatively small number of 52-week highs (lows) occur, the 52-week high (low) effect is significantly stronger.

3.7.4 Individual Trading and the 52-week High and Low Return Effects

Previous sections demonstrated that 52-week highs and lows predict significantly more positive buy-sell imbalances from sample individuals on the following day. I have also documented that 52-week highs and lows have a significant positive effect on next-day returns. While this evidence is suggestive of a causal relationship, it is certainly not definitive proof.

Unfortunately, establishing such proof is very difficult, given the nature and timing of the data studied herein. First, my data reports individual trades and returns at only the daily frequency. To definitely demonstrate that increased buying on day t causes abnormal returns on day t would require data with details at the intraday frequency. Without such data, there is no way to disentangle whether individual trading causes returns or whether individual purchases chase returns with a very short (i.e. intraday) lag.

Another shortcoming of the individual trade data is that it represents only a small sample of all individual traders. To the extent that this sample is representative of the entire population, it still provides excellent insight into the trading behavior of individuals, but its efficacy is somewhat limited to situations where trades can be aggregated over long periods of time or over many stocks.[29] It is far more difficult to make any inferences based on the trades of this small sample for individual stocks on single days. To illustrate this point, consider that of the over 4.7 million Nasdaq stock-day observations that occur during the individual trade database's sample period, only 252,424 (5.33%) experience trades by sample individuals. This reduction in sample size

[29] Indeed, such an aggregation was employed in Section 3.5.

is further complicated by the fact that 52-week highs and lows occur somewhat infrequently. In the current study, I observe only 13,325 52-week highs and 5,000 52-week lows that have corresponding individual trade data, as compared to a total of over 624,000 combined highs and lows for the full sample.

In this section I attempt to exploit these few observations for which trading does occur to strengthen the proposed relationship between individual trading and the return effects that follow 52-week highs and lows. First, the observations for which I observe at least one trade are divided into two groups: those with positive buy-sell imbalances (n=131,515) and those with negative buy-sell imbalances (n=118,632). I then run the Fama-Macbeth regressions described in Section 3.6.1 within each of these two groups.

The regression results in Table 3.8 demonstrate the relationship between individual trading and the 52-week high and low effects on returns. Note that on days when individuals are net sellers (BSI<0) following 52-week highs or lows, the coefficient estimates for the *High* and *Low* dummies are -65 and -13 basis points, respectively. On the other hand, when BSI is positive following a 52-week high or low, the coefficient estimates on the *High* and *Low* dummies are +13 and +55 basis points, respectively. Unfortunately, the dramatic drop in the number of observations prevents these coefficients from being estimated more precisely, as only one of the four attention proxies in the two regressions is significant at the 10 percent level.

To increase the power of the tests, in column three I use all observations for which we observe individual trades and include an interaction term between the attention dummies and a dummy variable indicating whether the observed buy-sell imbalance is positive. The positive and significant coefficient on the interaction between the 52-week high variable and the positive BSI dummy indicates that next-day returns are significantly larger at the one percent confidence level on days following 52-week highs that are accompanied by individual net buying. These findings suggest that only 52-week highs that generate attention, in the form of increased buying on the part of individuals on the day following the attention-grabbing event, experience excess positive returns on the day following the event.

3.8 ROBUSTNESS TESTS

This section contains two robustness tests for the main return predictability results presented in Table 3.4. In the first test, I divide the sample into two subperiods to ascertain whether the results are robust in the first and second halves of the sample. In the second test, I include five lags of each of the control variables to see what impact, if any, these lagged characteristics have on the results.

3.8.1 Subperiod Analysis

Recall that the sample analyzed in this paper consists of only Nasdaq stocks, due to the availability of closing bid and ask information for these securities on CRSP. This fact could be cause for concern, since the rise of the tech bubble of the late ‘90s and the subsequent crash make up a sizeable portion of the 20-year sample period. In this section, I recreate the regressions of Table 3.4 after dividing the sample into two subperiods, in order to ensure that the results for the full sample are not driven primarily by the anomalous price behavior of Nasdaq stocks during the latter portion of the sample. Table 3.9 reports the results from the subperiod regressions.

Columns I and IV of Table 3.9 indicate that the 52-week high and low effects are not driven by the tumultuous past 10-year period, as the positive impacts of 52-week highs and lows on subsequent returns are present in both subperiods. Note that the coefficient for the 52-week high variable is positive and significant in all six of the subperiod regressions presented in Table 3.9, while the coefficient on the 52-week low variable is positive in all regressions, and significant in all but the last column. Finally, it is interesting to observe that the relative strengths of the two attention variables appear to switch in the two subperiods; the 52-week low effect appears stronger during the earlier half, while the 52-week high effect is larger in the later half.

3.8.2 Additional Lags of Control Variables

Earlier results have shown that 52-week highs and lows predict positive next-day abnormal returns, even after controlling for event-day stock characteristics such as return, excess turnover, volume, liquidity and volatility. However, the analysis has not, to this point, taken into consideration the effects of these characteristics on days leading up to the event. Table 3.10 reports regression results after including the value of each control variable on each of days t-5 through t-1.

The results in Table 3.10 reinforce the importance of including these additional lags, particularly for past daily returns. Notice that the coefficient estimates for each of the five past return variables is positive and significant at the one percent confidence level. Note also that while returns exhibit strong positive autocorrelation with returns from the previous day, they are negatively correlated with returns on days t-2 through t-5.

Table 3.10 also reveals the impact that the additional control variables have on the attention proxies. A comparison to the final column of Table 3.4 reveals that the estimate for the 52-week high dummy increases from 14.6 basis points in Table 3.4 to 23.8 basis points in Table 3.10. Conversely, the coefficient estimate on the 52-week low dummy decreases from 13.8 basis points in Table 3.4 to 9.6 basis points in Table 3.10. This reduction in the estimate for the 52-week low coefficient also impacts its statistical significance, as zero becomes barely included in the coefficient's 10 percent confidence interval (p value = 0.1074).

3.9 ECONOMIC SIGNIFICANCE

While the preceding sections have established the statistical significance of the 52-week high and low effects, this section addresses the economic significance of the findings. Consider first the magnitude of the attention effects in isolation, represented by the coefficients on the 52-week high and low variables when all the control variables are included in the regression. Column VI of Table 3.4 indicates that, after controlling for the various event-day characteristics discussed in Section 3.6, the size of the 52-week high (low) effect on next-day stock return is (14.6) 13.8 basis points. While this effect may appear small at first, it is important to consider that the return horizon under consideration is only one day. Consider that the average daily return for all stocks in the sample is approximately 10 basis points.[30] Thus, the incidence of a 52-week high (low) causes returns to increase on the next day by 146% (138%) over the average daily return. Furthermore, the magnitudes of these effects compare favorably with other daily return effects that have been documented in the literature. For example, Hirshleifer and Shumway (2003) document that stocks earn average daily returns that are 8.8 basis points lower when it is overcast as opposed to sunny, and Tetlock (2006) finds that a one standard deviation increase in pessimism in the *Wall Street Journal* predicts a decrease in Dow Jones returns equal to 8.1 basis points over the next day.[31]

Another method of assessing the economic significance of the results is to consider the returns to a simple, theoretical trading strategy. Column I of Table 3.4 reveals that 52-week high (low) stocks earn average excess returns of 30.7 (9.4) basis points on the following day. Consider a trader who employs the following trading strategy. At the end of day t-1 (or beginning of day t), she purchases one randomly-selected stock that reached a new 52-week high on day t-1 and one randomly-selected stock that hit a 52-week low on day t-1. She then sells these stocks after holding them for a period of one day, at which time she purchases two stocks that hit 52-week highs and lows on day t. In a frictionless market, following this strategy will produce expected excess returns of 40.1 (30.7 + 9.4) basis points per day. Repeating this strategy daily, the

[30] This average daily return is found by averaging the one-day holding period returns, as reported by CRSP, overall all days and all stocks in the sample.

[31] The standardization of the variables in Tetlock's paper make it difficult to compare directly to the binary variables of interest in this study. Note that a change in Tetlock's pessimism variable from two standard deviations below the mean to two standard deviations above the mean would result in a 32.4 bp effect.

investor would expect to earn an annual excess return of approximately 180% per year.[32] Of course, this theoretical trading strategy relies on two unrealistic assumptions: zero trading fees and the ability to buy and sell at the closing bid-ask midpoint. In the presence of trading fees and bid-ask spreads, it is unlikely that the strategy, which requires daily transactions, will be profitable. Even if transaction costs do completely eliminate the profits, however, this example shows that the magnitudes of the 52-week high and low effects are not trivial.

Finally, it should be noted that if the 52-week high and low effects discussed in this paper do result from attention-driven buying, then the estimates in this paper are likely to understate the true size of the effects. While the measures in this paper only capture the return effect on the day following the event, the information that a 52-week high or low has occurred is actually revealed to the market at some time on the event day. If some attention-driven traders begin buying 52-week high and low stocks before market close on the event day, then the effect that these traders have on prices will not be captured.

3.10 CONCLUSIONS

Previous research has demonstrated that, consistent with the attention hypothesis, individuals are net buyers of stocks that capture their attention. In this study, I postulate that if this attention-driven buying by individuals is substantial, then it could exert significant upward price pressure on the securities in which these individuals trade following attention-grabbing events. Accordingly, I find strong evidence of predictably

[32] Calculated as $1.0041^{252} - 1$.

positive individual stock returns immediately following two events that capture the attention of investors: 52-week highs and lows.

Using daily Fama-MacBeth regressions over twenty years, I find that stocks experiencing 52-week highs and lows earn statistically and economically significant abnormal returns on days immediately following these events. After controlling for autocorrelation in the return data, the size of this outperformance is estimated to be 22.1 basis points for 52-week high stocks and 29.9 basis points for 52-week lows. The effect of a 52-week high or low on next-day returns remains positive and significant after controlling for various event-day, firm-level characteristics such as return, turnover, volatility and liquidity. Furthermore, the 52-week high effect on returns is shown to be robust to subsample analysis and to controlling for weekly seasonality in the control variables, although the statistical power of the 52-week low effect diminishes somewhat in these additional robustness tests.

In a further test of these results, I repeat the regressions after excluding from the sample a) the top 50% of all 52-week highs and lows as ranked on event-day absolute return, turnover or intraday volatility, and b) any 52-week highs or lows that occur within plus or minus three days of the underlying firm's earnings announcement. The positive and significant effect of 52-week highs and lows on future returns obtains even in this subsample of events which occur on relatively uneventful days, indicating that it is the occurrence of the event itself, and not any abnormal price or trading activity that may accompany it, which causes the effect.

Further tests offer strong support that the 52-week high and low effects do, in fact, result from the extra attention paid to these stocks. First, using individual household trade data from a large brokerage house, I document that individuals' buy-sell imbalances for stocks that hit a 52-week high or low on the previous day are significantly more positive than for all other stocks. While this result is consistent with the predictions of the attention hypothesis, other possible explanations for this behavior, such as trend chasing, contrarian trading or anchoring on 52-week highs (as proposed by George and Hwang (2004)), cannot simultaneously explain why investors would increase their net purchases following both 52-week highs and 52-week lows. Moreover, I show evidence that the predictable returns following 52-week highs are confined to those events that exhibit positive individual buy-sell imbalances on the following day.

Next, I document that on days when 52-week highs (lows) do the best job of focusing an investor's attention on a concentrated group of stocks, i.e. when a relatively small number of 52-week highs (lows) occurs, the 52-week high (low) effect is significantly stronger. Again, this result is consistent with the hypothesis that the abnormal returns following 52-week high and low events are the result of attention-based trading.

Finally, a powerful result emerges from comparing returns subsequent to 52-week highs and lows to returns following similarly-defined highs and lows over slightly shorter reference periods. In this experiment, I examine the effect that 47-, 49 and 51-week highs and lows, which are not simultaneously 52-week highs or lows, have on next-day returns. Interestingly, the tests reveal that these alternative high and low events have no significant impact on next-day returns. This result is particularly revealing, since the properties of the alternative high and low stocks should be almost identical to the properties of the 52-week high and low stocks; the only difference is that 52-week highs and lows are prominently followed and highly publicized, while 47-, 49- and 51-week highs and lows are not. Thus, these results strongly support the notion that the significant, positive effects of 52-week highs and lows on future returns are the direct result of the extra attention paid to these particular events.

Table 3.1: Yearly Summary Statistics

This table provides year-by-year summary statistics of the sample. Numbers reflect the average or median sample stock characteristics measured on the first trading day of each year. Size, price and closing bid and ask data come from the CRSP daily stock files, while analyst coverage is taken from I/B/E/S. The sample consists of stocks listed on NASDAQ during the period January 1986 to December 2005. To be included in the sample, a security must have data on closing bid, closing ask, and intraday high and low prices for at least 52 consecutive weeks. Stocks whose prices are below one dollar at the beginning of a given year are excluded from the sample in that year. *Size* is calculated as daily stock price multiplied by the number of shares outstanding and is reported in millions of dollars. *Firms* is the number of stocks included in the sample in a given year.

	Firms	Size (in millions)		Price		Bid-ask spread	
Year		Mean	Median	Mean	Median	Mean	Median
1986	2,485	79.87	25.57	13.39	8.50	0.47	0.25
1987	2,520	82.18	26.35	12.40	7.75	0.47	0.25
1988	2,782	78.86	22.90	10.32	6.75	0.56	0.38
1989	2,872	82.27	24.91	11.01	7.38	0.46	0.25
1990	2,751	100.89	26.05	11.60	7.44	0.49	0.25
1991	2,440	92.91	21.35	9.39	5.88	0.57	0.50
1992	2,597	145.06	29.08	11.81	7.25	0.61	0.50
1993	2,825	160.66	35.70	12.30	8.13	0.60	0.50
1994	3,164	183.75	46.08	12.92	9.00	0.57	0.50
1995	3,489	176.22	43.81	11.96	8.25	0.54	0.38
1996	3,770	245.97	54.85	14.05	9.88	0.54	0.38
1997	4,021	304.52	61.80	14.00	10.13	0.50	0.38
1998	4,328	371.47	73.05	15.37	10.50	0.44	0.25
1999	4,152	580.92	66.33	13.84	9.00	0.33	0.19
2000	4,048	1,094.72	77.10	17.30	9.50	0.30	0.19
2001	3,609	873.83	73.79	12.99	8.00	0.27	0.13
2002	3,541	807.65	95.83	13.23	9.13	0.20	0.10
2003	3,075	644.69	91.30	12.33	8.10	0.13	0.07
2004	3,156	949.96	158.04	16.62	12.47	0.15	0.06
2005	3,003	1,066.50	191.15	18.13	13.63	0.12	0.04

Table 3.2: Summary Statistics for 52-week High and Low Events

This table reports the number of observed 52-week highs and lows for each year of the sample, as well as the average event-day return associated with each. The table also reports the annual return on the value-weighted Nasdaq index for each year. A stock is said to have hit a 52-week high on any day when its highest intraday price exceeds the highest intraday price for the past 52 weeks for that stock. Similarly, a 52-week low occurs when a stock's intraday low price is smaller than the lowest intraday price experienced in the past 52 weeks. Returns are calculated using the bid-ask midpoint, the simple average of the closing bid and ask prices for a stock. The return is then defined as $return_t = (midpoint_t / midpoint_{t-1}) - 1$, where day t is the day on which a stock reaches a 52-week high or low. All returns are expressed as percentages.

	52-week Highs		52-week Lows		Nasdaq Index
Year	**Events**	**Avg Return**	**Events**	**Avg Return**	**Annual Return**
1986	13,768	2.05	9,794	-5.08	8.02
1987	9,213	3.80	16,416	-6.28	-3.98
1988	5,587	3.65	6,074	-5.33	17.88
1989	14,623	2.99	10,045	-4.82	21.22
1990	5,542	4.36	21,873	-4.75	-15.47
1991	13,770	3.85	4,546	-6.03	59.64
1992	14,293	4.16	6,920	-6.32	16.35
1993	17,979	4.48	6,617	-6.06	14.93
1994	10,622	4.11	14,378	-5.64	-3.21
1995	24,265	4.37	9,542	-6.28	40.26
1996	20,432	4.58	12,486	-6.27	22.23
1997	29,993	3.56	15,727	-6.50	22.23
1998	15,629	5.28	31,854	-6.22	39.31
1999	18,752	7.28	15,093	-5.25	83.67
2000	17,654	7.77	29,094	-6.58	-39.50
2001	16,061	3.93	20,486	-6.76	-20.79
2002	15,647	5.04	23,062	-5.67	-30.99
2003	36,858	3.31	4,825	-4.89	50.79
2004	24,086	3.52	10,058	-4.10	9.27
2005	17,968	3.44	13,230	-3.58	2.04

Table 3.3: Individual Trading Following 52-week Highs and Lows

This table reports the mean and standard deviation of daily individual buy-sell imbalances for stocks sorted into three groups: observations for which the previous day was a 52-week high, observations for which the previous day was a 52-week low, and all other observations. Following the methodology in Barber and Odean (2006), I define the daily buy-sell imbalance for each group as follows:

$$BSI_{pt} = \frac{\sum_{i=1}^{n_{pt}} NB_{it} - \sum_{i=1}^{n_{pt}} NS_{it}}{\sum_{i=1}^{n_{pt}} NB_{it} + \sum_{i=1}^{n_{pt}} NS_{it}}$$

where n_{pt} is the number of stocks in group p on day t, NB_{it} is the number of purchases of stock i on day t, and NS_{it} is the number of sales of stock i on day t. In Panel B, I calculate buy-sell imbalances based on the market value of trades by substituting the value of stock i bought (or sold) on day t for NB_{it} (or NS_{it}) in the equation above. The data contains information on the trades of 78,000 households with accounts at a large discount brokerage firm for the period January 1991 - November 1996. Trading data is aggregated over all sample households for each stock on each day. The row labeled High (Low) – Others reports the difference in the mean BSI of stocks in the Previous-day High (Low) group and the All Others group. The t-statistics for differences in means are in italics.

Panel A: Number of Trades

	Mean BSI	Std. Error
Previous-day High	11.54%	1.18%
Previous-day Low	37.54%	1.61%
All Others	5.42%	0.39%
High - Others	6.12%	
t-stat	*4.92*	
Low - Others	32.12%	
t-stat	*19.36*	

Panel B: Value of Trades

	Mean BSI	Std. Error
Previous-day High	8.72%	1.19%
Previous-day Low	38.95%	1.58%
All Others	0.93%	0.46%
High - Others	7.80%	
t-stat	*6.12*	
Low - Others	38.02%	
t-stat	*23.11*	

Table 3.4: Fama-MacBeth Regression Estimates

This table reports the Fama-MacBeth coefficient estimates and corresponding t-statistics from daily cross-sectional regressions of the form:

$$\textit{Market-adjusted ret}_{j,t} = \alpha + \beta_1 \cdot \textit{52-week high}_{j,t-1} + \beta_2 \cdot \textit{52-week low}_{j,t-1} + \gamma \cdot X_{j,t-1} + \varepsilon$$

where *52-week high*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1, *52-week low*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week low at time t-1, and $X_{j,t-1}$ is a vector of t-1 measurable, firm-specific control variables. Daily return is calculated as the percentage change in the bid-ask midpoint, which is defined as midpoint$_{j,t}$ = (closing ask$_{j,t}$ – closing bid$_{j,t}$) / 2. *Return*$_{j,t-1}$ is thus defined as (midpoint$_{j,t-1}$ / midpoint$_{j,t-2}$) -1. *Turnover*$_{j,t-1}$ is defined as volume$_{j,t-1}$ / shares outstanding$_{j,t-1}$. *Relative spread*$_{j,t-1}$ is defined as (closing ask$_{j,t-1}$ – closing bid$_{j,t-1}$) / midpoint$_{j,t-1}$. *Return*$^2_{j,t-1}$ is simply *return*$_{j,t-1}$ squared. *Abnormal turnover*$_{j,t-1}$ is defined on page 19, and the *Ret* VolumeDummy*$_{j,t-1}$ variables are explained on page 20. The dependent variable, *market-adjusted return*$_{t,j}$, is defined as the raw return, calculated from bid-ask midpoints, on day t minus the daily return on the Nasdaq value-weighted index on day t. All coefficient estimates are reported in basis points. Newey-West corrected t-statistics are reported in italics.

Dependent variable: Market-adjusted return$_{j,t}$						
	(I)	(II)	(III)	(IV)	(V)	(VI)
Intercept	-7.64	-7.57	-9.65	-7.23	-9.61	-9.59
	-3.28	*-3.21*	*-3.99*	*-3.07*	*-3.88*	*-4.04*
52-week High$_{j,t-1}$	30.69	22.09	17.31	15.50	15.26	14.60
	15.44	*6.18*	*5.62*	*4.74*	*7.40*	*4.20*
52-week Low$_{j,t-1}$	9.45	29.89	25.48	24.85	26.15	13.76
	2.97	*7.34*	*5.83*	*5.79*	*7.21*	*2.67*
Return$_{j,t-1}$		290.84	281.75	285.66		317.30
		4.53	*4.22*	*4.32*		*4.35*
Turnover$_{j,t-1}$			541.75		510.70	278.80
			10.37		*13.49*	*4.31*
Abnormal Turnover$_{j,t-1}$				665.05		
				18.13		
Ret*HighVolumeDummy$_{j,t-1}$					315.19	
					16.49	
Ret*LowVolumeDummy$_{j,t-1}$					309.83	
					1.75	
Relative Spread$_{j,t-1}$						63.30
						1.01
Return$^2_{j,t-1}$						1518.72
						5.76
Avg. Adjusted R^2	0.0021	0.0083	0.0121	0.0104	0.0139	0.0200

Table 3.5: Estimates for Highs and Lows with Different Reference Periods

This table reports the Fama-MacBeth coefficient estimates and corresponding t-statistics from daily cross-sectional regressions of the form:

$$\textit{Market-adjusted ret}_{j,t} = \alpha + \beta_1 \cdot \textit{52-week high}_{j,t-1} + \beta_2 \cdot \textit{52-week low}_{j,t-1} + \gamma \cdot X_{j,t-1} + \varepsilon$$

where *52-week high*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1, *52-week low*$_{j,t-1}$ takes the value of one if stock j hit a 52-week low at time t-1, and $X_{j,t-1}$ is a vector of t-1 measurable, firm-specific control variables. In Columns II-IV, the 52-week dummy variables are replaced with *n-week high* and *n-week low* (n=47, 49, 51). *N-week high*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit an n-week high, but not a 52-week high, on day t-1, and *n-week low*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit an n-week low, but not a 52-week low, on day t-1. Coefficients are reported in basis points. Newey-West corrected t-statistics are reported in italics.

Dependent variable: Market-adjusted return$_{j,t}$				
	(I)	(II)	(III)	(IV)
Intercept	-9.59	-9.37	-9.36	-9.34
	-4.04	*-3.83*	*-3.82*	*-3.82*
52-week High$_{j,t-1}$	14.60			
	4.20			
52-week Low$_{j,t-1}$	13.76			
	2.67			
47-week High$_{j,t-1}$		-8.18		
		-1.06		
47-week Low$_{j,t-1}$		7.52		
		1.58		
49-week High$_{j,t-1}$			-9.70	
			-1.24	
49-week Low$_{j,t-1}$			5.79	
			1.29	
51-week High$_{j,t-1}$				-7.25
				-1.02
51-week Low$_{j,t-1}$				4.32
				1.02
Return$_{j,t-1}$	317.30	315.64	315.61	315.94
	4.35	*4.53*	*4.53*	*4.56*
Turnover$_{j,t-1}$	278.80	294.31	295.56	294.77
	4.31	*4.58*	*4.60*	*4.60*
Relative Spread$_{j,t-1}$	63.30	63.40	63.42	63.33
	1.01	*1.03*	*1.03*	*1.03*
Return$^2_{j,t-1}$	1518.72	1555.16	1557.97	1557.49
	5.76	*5.88*	*5.89*	*5.89*
Avg. Adjusted R^2	0.0200	0.0190	0.0189	0.0187

Table 3.6: Regression Estimates after Excluding "Informative" Highs and Lows

This table reports the Fama-MacBeth coefficient estimates and corresponding t-statistics for the subsample resulting from excluding all 52-week highs and lows that occur within +/- 3 days of the underlying firm's earnings announcement, as well as any 52-week highs (lows) which rank in the top 50% of all 52-week highs (lows) when sorted on event-day absolute return, turnover, or volatility. On this subsample, I run daily cross-sectional regressions of the form:

$$\textit{Market-adjusted ret}_{j,t} = \alpha + \beta_1 \cdot \textit{52-week high}_{j,t-1} + \beta_2 \cdot \textit{52-week low}_{j,t-1} + \gamma \cdot X_{j,t-1} + \varepsilon$$

where *52-week high*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1, *52-week low*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week low at time t-1, and $X_{j,t-1}$ is a vector of t-1 measurable, firm-specific control variables. Daily return is calculated as the percentage change in the bid-ask midpoint, which is defined as $\text{midpoint}_{j,t}$ = (closing $\text{ask}_{j,t}$ – closing $\text{bid}_{j,t}$) / 2. *Return*$_{j,t-1}$ is thus defined as ($\text{midpoint}_{j,t-1}$ / $\text{midpoint}_{j,t-2}$) -1. *Turnover*$_{j,t-1}$ is defined as $\text{volume}_{j,t-1}$ / shares $\text{outstanding}_{j,t-1}$. *Relative spread*$_{j,t-1}$ is defined as (closing $\text{ask}_{j,t-1}$ – closing $\text{bid}_{j,t-1}$) / $\text{midpoint}_{j,t-1}$. *Return*$^2_{j,t-1}$ is simply *return*$_{j,t-1}$ squared. The dependent variable, *market-adjusted return*$_{t,j}$, is defined as the raw return, calculated from bid-ask midpoints, on day t minus the daily return on the Nasdaq value-weighted index on day t. All coefficient estimates are reported in basis points. Newey-West corrected t-statistics are reported in italics.

Dependent variable: Market-adjusted $\text{return}_{j,t}$	
Intercept	-9.57
	-3.02
52-week $\text{High}_{j,t-1}$	21.45
	5.91
52-week $\text{Low}_{j,t-1}$	17.86
	4.71
$\text{Return}_{j,t-1}$	409.15
	3.94
$\text{Turnover}_{j,t-1}$	330.25
	3.60
Relative $\text{Spread}_{j,t-1}$	51.28
	0.61
$\text{Return}^2_{j,t-1}$	1581.28
	4.37
Avg. Adjusted R^2	0.018

Table 3.7: Coefficient Estimates within Deciles Sorted on Daily Number of High or Low Events

This table reports the Fama-MacBeth coefficient estimates of β_1 (Panel A) and β_2 (Panel B) from the regression:

$$Mkt\text{-}adj\ ret_{j,t} = \alpha + \beta_1 \cdot 52\text{-}wk\ high_{j,t-1} + \beta_2 \cdot 52\text{-}wk\ low_{j,t-1} + \beta_3 \cdot ret_{j,t-1} + \beta_4 \cdot turnover_{j,t-1} + \beta_5 \cdot rspread_{j,t-1} + \beta_6 \cdot ret^2_{j,t-1} + \varepsilon$$

within each of 10 deciles. For Panel A, all sample days are first sorted into deciles by the total number of 52-week highs that occurred on the previous day. For Panel B, sample days are sorted into deciles by the total number of 52-week lows that occurred on the previous day. Variable definitions are as follows. *52-wk* $high_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1. *52-wk* $low_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week low at time t-1. Daily return is calculated as the percentage change in the bid-ask midpoint, which is defined as $midpoint_{j,t}$ = (closing $ask_{j,t}$ – closing $bid_{j,t}$) / 2. $Ret_{j,t-1}$ is thus defined as ($midpoint_{j,t-1}$ / $midpoint_{j,t-2}$) -1. $Turnover_{j,t-1}$ is defined as $volume_{j,t-1}$ / shares $outstanding_{j,t-1}$. $Rspread_{j,t-1}$ is defined as (closing $ask_{j,t-1}$ – closing $bid_{j,t-1}$) / $midpoint_{j,t-1}$. $Ret^2_{j,t-1}$ is simply $ret_{j,t-1}$ squared. The dependent variable, *mkt-adj* $ret_{t,j}$, is defined as the raw return, calculated from bid-ask midpoints, on day t minus the daily return on the Nasdaq value-weighted index on day t. All coefficient estimates are reported in basis points. Statistical significance at the 10, 5, and 1 percent confidence levels is denoted by *, **, and ***, respectively.

Panel A: Sorted by daily # of 52-week highs		Panel B: Sorted by daily # of 52-week lows	
Decile	52-week High Coeff. Estimate	Decile	52-week Low Coeff. Estimate
1 (fewest)	25.10 **	1 (fewest)	34.78 **
2	17.05 ***	2	29.58 ***
3	15.27 ***	3	16.76 **
4	11.00 **	4	6.28
5	28.02 **	5	-21.05
6	18.15 ***	6	18.79 **
7	14.23 ***	7	19.60 ***
8	6.67 **	8	20.25 **
9	4.52	9	16.07 **
10 (most)	4.58	10 (most)	-0.75
1 minus 10	20.51 *	1 minus 10	35.54 **

Table 3.8: Return Effects and Contemporaneous Individual Trading

This table reports the Fama-MacBeth coefficient estimates from the regression:

$$Mkt\text{-}adj\ ret_{j,t} = \alpha + \beta_1 \cdot 52\text{-}wk\ high_{j,t-1} + \beta_2 \cdot 52\text{-}wk\ low_{j,t-1} + \beta_3 \cdot ret_{j,t-1} + \beta_4 \cdot turnover_{j,t-1} + \beta_5 \cdot rspread_{j,t-1} + \beta_6 \cdot ret^2_{j,t-1} + \varepsilon.$$

Column one reports results for stocks with negative buy sell imbalances (BSI) on day t, column two reports results for stocks with positive BSI on day t, and column three uses all observations for which BSI is observed. $BSI_{j,t}$ is defined as the total number of buys of stock j minus the total number of sells by sample individuals on day t. *52-wk high*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1. *52-wk low*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week low at time t-1. *BSI+*$_{j,t}$ is a dummy that takes a value of 1 if stock j has positive BSI on day t and is interacted with all of the variables from the first column. Daily return is calculated as the percentage change in the bid-ask midpoint, which is defined as $midpoint_{j,t}$ = (closing $ask_{j,t}$ – closing $bid_{j,t}$) / 2. *Ret*$_{j,t-1}$ is thus defined as ($midpoint_{j,t-1}$ / $midpoint_{j,t-2}$) -1. *Turnover*$_{j,t-1}$ is defined as $volume_{j,t-1}$ / shares $outstanding_{j,t-1}$. *Rspread*$_{j,t-1}$ is defined as (closing $ask_{j,t-1}$ – closing $bid_{j,t-1}$) / $midpoint_{j,t-1}$. $Ret^2_{j,t-1}$ is simply $ret_{j,t-1}$ squared. The dependent variable, *mkt-adj ret*$_{t,j}$, is defined as the raw return, calculated from bid-ask midpoints, on day t minus the daily return on the Nasdaq value-weighted index on day t. All coefficient estimates are reported in basis points. Newey-West corrected t-statistics are reported in italics.

Dependent variable: Market-adjusted $return_{j,t}$			
	BSI_t<0	BSI_t>0	All
Intercept	20.69	-50.30	-17.40
	4.75	*-12.04*	*-5.10*
52-week $High_{j,t-1}$	-65.40	13.39	-73.00
	-1.80	*1.20*	*-1.89*
52-week $Low_{j,t-1}$	-13.00	54.67	57.54
	-0.37	*0.92*	*0.88*
52-week $High_{j,t-1}$* $BSI+_{j,t}$			47.18
			2.80
52-week $Low_{j,t-1}$* $BSI+_{j,t}$			7.48
			0.19
$Return_{j,t-1}$	1891.98	1505.90	1878.38
	23.13	*21.81*	*23.15*
$Turnover_{j,t-1}$	445.04	389.82	1409.12
	2.40	*2.52*	*8.09*
Relative $Spread_{j,t-1}$	816.38	189.71	1397.68
	6.00	*1.72*	*12.01*
$Return^2_{j,t-1}$	-2544.00	3599.20	-1230.70
	-1.97	*5.53*	*-2.50*
Avg. Adjusted R^2	0.093	0.090	0.097

Table 3.9: Subperiod Regression Estimates

This table reports the Fama-MacBeth coefficient estimates and corresponding t-statistics from daily cross-sectional regressions of the form:

$$\textit{Market-adjusted } ret_{j,t} = \alpha + \beta_1 \cdot 52\text{-}week\ high_{j,t-1} + \beta_2 \cdot 52\text{-}week\ low_{j,t-1} + \gamma \cdot X_{j,t-1} + \varepsilon$$

where *52-week* $high_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1, *52-week* $low_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week low at time t-1, and $X_{j,t-1}$ is a vector of t-1 measurable, firm-specific control variables. Daily return is calculated as the percentage change in the bid-ask midpoint, which is defined as $midpoint_{j,t}$ = (closing $ask_{j,t}$ – closing $bid_{j,t}$) / 2. $Return_{j,t-1}$ is thus defined as ($midpoint_{j,t-1}$ / $midpoint_{j,t-2}$) -1. $Turnover_{j,t-1}$ is defined as $volume_{j,t-1}$ / shares $outstanding_{j,t-1}$. *Relative* $spread_{j,t-1}$ is defined as (closing $ask_{j,t-1}$ – closing $bid_{j,t-1}$) / $midpoint_{j,t-1}$. $Return^2_{j,t-1}$ is simply $return_{j,t-1}$ squared. The dependent variable, *market-adjusted* $return_{t,j}$, is defined as the raw return, calculated from bid-ask midpoints, on day t minus the daily return on the Nasdaq value-weighted index on day t. All coefficient estimates are reported in basis points. Newey-West corrected t-statistics are reported in italics.

Dependent variable: Market-adjusted $return_{j,t}$						
	Sample period					
	1986-1995			1996-2005		
	(I)	(II)	(III)	(IV)	(V)	(VI)
Intercept	-12.84	-15.65	-13.11	-2.32	-3.67	-6.09
	-3.10	*-3.75*	*-3.18*	*-1.04*	*-1.52*	*-2.57*
52-week $High_{j,t-1}$	17.81	11.02	6.02	26.35	23.57	23.14
	5.89	*3.75*	*1.97*	*4.08*	*4.37*	*3.73*
52-week $Low_{j,t-1}$	38.07	32.66	20.55	21.74	18.33	7.00
	8.90	*7.67*	*4.73*	*3.15*	*2.41*	*0.75*
$Return_{j,t-1}$	745.40	736.34	845.02	-161.91	-171.04	-208.31
	40.26	*40.54*	*44.23*	*-1.31*	*-1.33*	*-1.48*
$Turnover_{j,t-1}$		811.56	557.03		273.01	1.67
		16.18	*12.13*		*3.03*	*0.01*
Relative $Spread_{j,t-1}$			-41.27			167.46
			-3.95			*2.50*
$Return^2_{j,t-1}$			1748.89			1289.47
			16.01			*1.35*
Avg. Adjusted R^2	0.011	0.014	0.023	0.006	0.010	0.017

Table 3.10: Regression Estimates with Additional Lags of Control Variables

This table reports the Fama-MacBeth coefficient estimates and corresponding t-statistics from daily cross-sectional regressions of the form:

$$\textit{Market-adjusted ret}_{j,t} = \alpha + \beta_1 \cdot \textit{52-week high}_{j,t-1} + \beta_2 \cdot \textit{52-week low}_{j,t-1} + \sum_{n=1}^{5} \gamma_n \cdot X_{j,t-n} + \varepsilon$$

where *52-week high*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week high at time t-1, *52-week low*$_{j,t-1}$ is a dummy variable that takes the value of one if stock j hit a 52-week low at time t-1, and $X_{j,t-n}$ is a vector of t-n measurable, firm-specific control variables. Daily return is calculated as the percentage change in the bid-ask midpoint, which is defined as midpoint$_{j,t}$ = (closing ask$_{j,t}$ – closing bid$_{j,t}$) / 2. *Return*$_{j,t-1}$ is thus defined as (midpoint$_{j,t-1}$ / midpoint$_{j,t-2}$) -1. *Turnover*$_{j,t-1}$ is defined as volume$_{j,t-1}$ / shares outstanding$_{j,t-1}$. *Relative spread*$_{j,t-1}$ is defined as (closing ask$_{j,t-1}$ – closing bid$_{j,t-1}$) / midpoint$_{j,t-1}$. *Return*$^2_{j,t-1}$ is simply *return*$_{j,t-1}$ squared. The dependent variable, *market-adjusted return*$_{t,j}$, is defined as the raw return, calculated from bid-ask midpoints, on day t minus the daily return on the Nasdaq value-weighted index on day t. All coefficient estimates are reported in basis points. Newey-West corrected t-statistics are reported in italics.

Dependent variable: Market-adjusted return$_{j,t}$				
	Coefficient	Standard error	t-stat	P value
Intercept	-1.1	0.0003	-4.11	<.0001
52-week High$_{j,t-1}$	23.8	0.0004	6.74	<.0001
52-week Low$_{j,t-1}$	9.6	0.0006	1.61	0.1074
Return				
lag t-1	282.5	0.0068	4.17	<.0001
lag t-2	-146.2	0.0015	-9.81	<.0001
lag t-3	-149.5	0.0010	-14.61	<.0001
lag t-4	-93.1	0.0011	-8.87	<.0001
lag t-5	-67.0	0.0010	-7.03	<.0001
Turnover				
lag t-1	381.0	0.0093	4.11	<.0001
lag t-2	-60.9	0.0031	-1.98	0.0472
lag t-3	22.2	0.0036	0.62	0.5364
lag t-4	-28.5	0.0032	-0.90	0.3685
lag t-5	-16.0	0.0026	-0.61	0.5415
Relative Spread				
lag t-1	-193.3	0.0039	-5.03	<.0001
lag t-2	130.8	0.0055	2.38	0.0176
lag t-3	91.1	0.0043	2.12	0.0338
lag t-4	29.6	0.0018	1.63	0.1036
lag t-5	74.8	0.0019	3.93	<.0001
Return^2				
lag t-1	1495.7	0.0267	5.59	<.0001
lag t-2	175.7	0.0126	1.39	0.1634
lag t-3	16.9	0.0085	0.20	0.842
lag t-4	75.3	0.0055	1.38	0.1687
lag t-5	-50.3	0.0044	-1.15	0.2521
Avg. R^2	0.047			

CHAPTER 4: REPUTATION AND MUTUAL FUND CHOICE

4.1 INTRODUCTION

Reputational capital is an important asset for many firms, particularly for firms that offer experience goods because reputation provides a signal of product quality. A firm's reputation can be gained through many avenues such as marketing, performance, service, media coverage, and word-of-mouth. Reputation can influence the relation between lenders and borrowers (Diamond, 1989, 1991); can affect the actions of traders (Battalio, Ellul and Jennings, 2006) and financial analysts (Stickel, 1992; Jackson, 2005; Fang and Yasuda, 2006, 2007; and Ljungqvist, et al, 2006); can allow firms to charge premium prices for high quality products (Klein and Leffler, 1981; Allen, 1984; Milgrom and Roberts, 1986); and can affect pricing of IPOs (Carter and Manaster, 1990) among other aspects.

Mutual funds are a prime example of a product, an experience good, for which reputation should be an important attribute for the investor's purchase decision. Survey evidence suggests that this hypothesis is valid. Surveys of individual mutual fund shareholders have found that reputation was among the most important characteristics in the mutual fund shareholders' purchase decision. One survey found that fund manager reputation was the second most important factor (after investment performance track record) and a second survey found that fund family reputation was the third most important piece of information considered before the most recent fund purchase (after risk level and total return) (Capon, Fitzsimmons, Prince, 1996; ICI, 1997).[33] A survey of financial advisers found that fund manager reputation was the fifth most important characteristic after relative performance, fund objective, fund risk, and fund manager tenure (Jones, Lesseig and Smythe, 2005).

Although this survey evidence establishes that investors assert reputation as being important in their decisions, there has not been an examination of the effects of reputation on the mutual fund purchase decision. I conduct such an examination in this chapter through use of an ideal database for such an investigation: individual investors' mutual fund holdings and trade data from a large discount brokerage firm. By examining the investors' revealed choices of funds, I am able to divorce such decisions from service quality, because these investors are buying the funds through omnibus accounts and thus are not getting direct service from the fund advisor. Furthermore, I am able to abstract such decisions from the decision to stay within a particular fund family for ease of switching funds (e.g., Massa, 2003), for lessening of search costs or recordkeeping (Elton, Gruber, and Green, 2007) or due to restrictions caused by employee retirement plans (Elton, Gruber and Blake, 2006).

For the purposes of this study, I define reputation as the information, whether hard or soft, that together with past performance comprises an investor's expectations about the managerial competence of a mutual fund or mutual fund family.[34] My central

[33] It is not clear whether survey takers would interpret fund manager reputation to mean the reputation of the portfolio manager or the fund management company. In some sense mutual fund families could correspond to Tirole's (1996) concept of collective reputation, where group reputation is an aggregate of individual reputations.
[34] The term "information" may be somewhat of a misnomer, as we make no assumptions about the truthfulness or accuracy of the information that the investor uses to formulate her beliefs.

hypothesis is that if individual investors use fund family reputation as an important determinant of their mutual fund choices, then individuals should be more likely to concentrate their mutual fund holdings within the select family or families that they believe to be the most competent. This result should hold even if the individual is purchasing the mutual fund on a supermarket and thus does not have other incentives to stay within the same fund family. I find evidence to support this hypothesis, as individual investors in the supermarket sample are shown to cluster their mutual fund holdings within families.

Another notion that develops from the reputation literature is that an established reputation causes an investor's beliefs about the competence of the firm to dissipate slowly (Mailath and Samuelson, 2001). Given that the effects of reputation change slowly over time, I hypothesize that investors who value reputation should be more likely to be repeat customers of fund families.[35] Consistent with this hypothesis, I find that investors who have previously held funds from a particular fund family are significantly more likely to choose funds from the same family when they enter the marketplace for mutual funds again. My findings reveal that this previous family ownership effect is highly statistically and economically significant to individuals' mutual fund choices and is robust to the inclusion of a host of control variables that are known to affect investor choice. Indeed, I find that a fund is six times more likely to be chosen if it belongs to a family with which the individual has prior experience. Moreover, I find that this previous

[35] Milgrom and Roberts (1986) reinforce this hypothesis by noting the importance of facilitating repeat customers to the choice of firms who attempt to build reputation for their experience goods through uninformative advertising.

family ownership effect exists regardless of the quality of the return performance that the investor experienced previously with the family. This finding is consistent with the idea that reputation, in addition to past performance, plays an important role in individuals' mutual fund choices.

Next, I test whether the documented propensity of individuals to purchase repeatedly from the same family exists for the subsample of initiations into funds having objectives with which the investor has no prior investment experience. The motivation for such a test is straightforward. If the sizable effect of previous family ownership on individual choice diminishes in this subsample, then it is likely that an alignment in style preference between the investor and the fund family is driving the result. Alternatively, such a finding could also be consistent with the idea that families generate familiarity with investors by reporting member funds with similar investment objectives in the same prospectus. However, I find strong evidence that neither of these alternative explanations is responsible for the strong positive effect that previous experience with a family has on the probability that a fund from the same family is chosen in the future. In fact, the effect of the previous family ownership variable is marginally stronger among this subsample of initiations in new objective classes.

In a final test of the effect of fund family affiliation on investor choice, I examine whether the response of fund flows for individual investors to past returns is significantly different for funds belonging to families with stronger reputations, as proxied by family size and family age. Ippolioto (1992), among others, demonstrates that investors rely on past performance as a signal of mutual fund quality. Hence, the sensitivity of flow to past returns is a measure of how investors update their beliefs about funds. Mailath and Samuelson (2001) show that a strong reputation (i.e. when consumers have strong posterior expectations that the firm is competent) will result in consumers' beliefs changing only gradually as they receive additional signals. In the context of the current study, this gives rise to the hypothesis that investors will be slower to update their beliefs about mutual funds belonging to families with more established reputations. I find evidence to support this hypothesis, as flows into funds belonging to larger or older fund families are shown to be less sensitive to return performance than are fund flows for funds of smaller or younger families.

An important aspect of the results contained in this chapter is that they provide an additional rationale for organization in the mutual fund industry. Beyond economies of scale, the ability to lower investors' effective fees, or increasing the likelihood of producing a "star" fund and the accompanying spillover effects (Nanda, Wang and Zheng, 2004), mutual fund advisors have an incentive to organize into families because of the reputation synergies that develop from such an organization style. Such considerations appear to increase repeat customers and allow funds of more established families to compete on a dimension other than returns.

In addition to the reputation literature cited above, my research is related to several papers on the mutual fund selection decision such as Sirri and Tufano (1998), Del Guercio and Tkac (2002), and Huang, Wei and Yan (2006), and to papers on the organization of and decisions of mutual fund families such as Massa (2001, 2003), Khorana and Servaes (2003), and Gallaher, Kaniel and Starks (2006). Two papers closely related to my study in that they use the same sample of individual investors to examine mutual fund purchase and sell decisions are Barber, Odean and Zheng (2005) and Ivkovic and Weisbenner (2006); however, the focus of these papers is quite different from mine. Barber, Odean and Zheng employ the data to examine the effects of fees on mutual fund purchase decisions, and Ivkovic and Weisbenner employ the data to examine the existence of the disposition effect in mutual fund trades. Finally, Johnson (2004) analyzes a proprietary database of individual shareholders in one no-load mutual fund family, but his focus in on the predictability and costs related to short- and long-horizon investors.

The remainder of the chapter is organized as follows. The next section summarizes the data used in the analysis. Section 4.3 examines the extent to which investors concentrate their investments at the family level. Section 4.4 provides regression analysis of the effect of prior family ownership on the probability that a family's fund is selected. Section 4.5 provides additional tests to determine whether the results are consistent with a reputation explanation. Section 4.6 examines the effect of family reputation on flow-performance sensitivity. Section 4.7 concludes.

4.2 DATA DESCRIPTION

4.2.1 Data Sources

I employ data sets from two sources. First, I use a database of mutual fund holdings and trades from a large discount brokerage firm. This database contains information for 128,000 individual accounts from January 1991 through November

1996.[36] Second, I use the CRSP Mutual Fund Database (MFDB). I cross-reference the cusips in the discount brokerage dataset with the MFDB to identify any holdings or trades involving mutual funds during the sample period. For each mutual fund transaction, I identify the account number, the date of the trade, the number of shares of the fund bought or sold and the price at which the trade is executed.

From the MFDB, I collect information on the individual mutual funds. I obtain annual data on each mutual fund's net asset value (NAV), total net assets (TNA), front-end load, rear load, expense ratio, turnover ratio, the date the fund was first offered, the ICDI fund objective, and the fund's management company name and code.[37] I also obtain fund returns from the MFDB at monthly frequencies.

From these fund-level variables, I also construct family-level variables. Family TNA is the sum of the TNA for every fund belonging to the management company in each period; family objectives is the number of unique ICDI objectives represented by all family funds; and family age is the age of the oldest fund in each family as of the current period, rounded to the nearest year. All other family-level variables are computed as the weighted average of the fund-level variables of member funds, where funds are weighted by their TNA at the end of the prior period.

[36] This data set is described in detail in Barber and Odean (2000, 2001, and 2002). I thank Terry Odean for the use of this database.
[37] Information on management company and ICDI objective code is not available from CRSP until 1992. Thus, for years prior to 1992, I backfill these data items using funds' 1992 information.

4.2.2 Summary Statistics

Table 4.1 summarizes the annual mutual fund holdings of sample individual household accounts. Panel A reports the total number of households for each year in the database and their average mutual fund holdings. The table clearly displays the growing popularity of mutual funds among the sample individuals throughout the sample period. In January 1991, only 15.7 percent of the households hold any mutual funds in their brokerage accounts. By 1996, that figure more than doubles to 35.0 percent.[38] Consistent with the increase in the percentage of brokerage accounts holding mutual funds, the average percentage of total brokerage account assets held in mutual funds also increases every year in the sample.

Panel B reports mutual fund portfolio statistics for those households that held at least one mutual fund in a given year. The table shows that for such accounts, mutual funds make up a sizable portion of the investor's total account holdings. In fact, among the households that hold mutual funds, the average investor held over 45 percent of all brokerage assets in mutual funds in each year of the sample. By January of 1996, household accounts that invest in mutual funds have an average of 3.2 unique mutual funds worth an average (median) market value of $49,553 ($16,011).

Panel C of Table 4.1 provides the distribution of the average market value of mutual funds held in the discount brokerage accounts for those households holding at least one mutual fund. The distribution further exhibits the growth in mutual fund

[38] This is consistent with overall U.S. households. The Investment Company Institute (2006) reports that 36.8% of households in general held mutual funds in 1996.

holdings of sample individuals. While the median household holds only $6,905 in mutual funds at the beginning of the sample, that amount increases to $16,011 by 1996. For those individuals that invest the most in mutual funds, the growth is even more pronounced. Indeed, by 1996 ten percent of the mutual fund-holding households have more than $108,000 invested in mutual funds, and five percent have more than $194,000.

Table 4.2 reports annual summary statistics for the mutual funds in the CRSP database over the 1990-1996 sample period. Panel A contains summary statistics for all mutual funds found in the CRSP database, while Panel B summarizes the statistics for those mutual funds that were held at least once by the individuals in the discount brokerage database.[39] Consistent with the results in Table 4.1, Panel A demonstrates the explosive growth of the mutual fund industry during the sample period. Over the seven years, the number of share classes increases by almost 300 percent, from 3,684 to 10,125.[40]

Panel B demonstrates that the sample individual investors held only a small portion of the total mutual fund share classes that existed during the sample period. For instance, sample individual households collectively held only 770 of the 10,125 mutual fund share classes available to them in 1996. That is, a comparison of Panels A and B shows that while many funds were being added to the marketplace, the investors in the brokerage database were concentrated in a smaller set of funds. The number of funds in which these investors purchase shares increases by only 144 percent. Also note that the

[39] The CRSP database classifies each share class as a mutual fund. For ease of exposition through much of the paper, I simply term these "funds."

[40] See Nanda, Wang and Zheng (2005) for a discussion of share classes.

average Total Net Assets (TNA) for all funds in Panel A does not change materially over the years due to the addition of new funds each year, which keeps the average TNA low. In contrast, in Panel B the average TNA increases by almost 300 percent as the sample investors tend to invest in a smaller set of funds which are growing in assets under management. A comparison of Panels A and B also suggests that the sample individuals are drawn primarily to funds that are a good deal larger and older than the average fund. For example, the average fund held by a sample individual household in 1996 had total net assets of $1.58 billion and was approximately 15 years old. By contrast, the average TNA and age for all funds in 1996 was only $332 million and 6.6 years, respectively.

Using monthly fund returns from the CRSP MFDB, I calculate the 12-month buy-and-hold return for each fund in each year. I also compute a 12-month objective-adjusted return for each fund. To compute the objective-adjusted return variable, I first calculate the simple average of the 12-month compound return of all funds within each ICDI objective category in each year. A fund's objective-adjusted return is then defined as the fund's 12-month compound return minus the average compound return for all funds in its objective category. Comparing Panels A and B of Table 4.2, the returns on the funds in which the brokerage firm investors put their money tend to be higher on both a raw and objective-adjusted basis.[41]

Because my central interest is in the reputation of the fund family, I also aggregate the data to construct the mutual fund characteristic variables at the fund family

[41] The higher returns in Panel B are not necessarily evidence that investors have any selection ability. Since funds need only be held by one account at any time during the sample to be included in Panel B, the higher returns could reflect the tendency of investors to flock to funds with superior past performance. This behavior is demonstrated by the positive, significant coefficient on 12-month objective-adjusted return in Table 4.10. I address the issue of portfolio performance in Section 4.6.

level. Family TNA is defined as the sum of the TNA of each member fund belonging to the fund family in any year. Family age is defined as the maximum age among the funds comprising the family. Family-level total expenses, return and objective adjusted return are defined as weighted averages of the corresponding variables for member funds in each year, where member funds are weighted within each family by their TNA.

Panel A of Table 4.3 provides the averages for the family variables described above, along with the number of funds and number of objectives offered by the family, for each year of the sample. The table clearly demonstrates a growing trend throughout the sample period, as fund families increase not only in number, but also in average TNA, number of funds, and number of objectives offered. Table 4.3 also demonstrates an upward trend in average expenses during the sample, which may be caused by the addition of more complex objective classes with higher accompanied fees.[42] Panel B of Table 4.3 provides summary statistics for only those families held by at least one individual investor during the sample. Comparing Panels A and B reveals a strong preference by individual investors for older, larger, and broader fund families.

4.3 REPUTATION AND FUND FAMILY CONCENTRATION

If reputation is an important attribute to an individual's choice of mutual fund investment, then we should observe individuals investing in funds from the same fund family. To test this primary hypothesis, I first calculate the concentration of the sample individuals' mutual fund investments within the same families. I divide sample households by the number of different mutual funds that they hold at any time during the

[42] See, for example, the 2001 SEC Report on Mutual Fund Fees.

sample period. I then calculate the percentage of households within each group that holds at least *n* funds in the same mutual fund family. I also calculate the percentage of households within each group that concentrate all of their mutual fund holdings within the same family.

Table 4.4 reports the results from these tests. I find that for accounts with two or more funds, a substantial percentage are from the same family. For instance, 32 percent of investors who hold exactly two mutual funds over the entire sample choose both funds from the same family. Moreover, over half of individuals who hold three or more mutual funds choose at least two funds from the same family. Additionally, for individuals investing in a very large number of funds (more than ten), over 60% have more than four funds from the same family. The extent of this same-family concentration is surprisingly high, especially considering that Panel B of Table 4.3 shows that investors have a minimum of 143 families available to them at any time during the sample. A simple back-of-the-envelope calculation reveals that the probability of two randomly-selected funds from the marketplace belonging to the same family would be around 0.7 percent.[43] These results demonstrate that sample individuals pay particular attention to fund family membership when selecting their mutual fund investments.

[43] Calculated as (1/143). This estimation relies on the obviously erroneous simplifying assumption that all of the 143 families contain the same number of funds. This is also a conservative estimate, counting only the fund families held by sample individuals in 1990. The other extreme would be to use the 457 families on the CRSP database in 1996, which would be 0.22 percent.

4.4 PREVIOUS FAMILY OWNERSHIP AND FUND CHOICE

4.4.1 Baseline Regression Analysis

The benefit of reputation for experience goods such as mutual funds is that consumers will repeat purchases of the same product (Milgrom and Roberts, 1986). This implies that if reputation is an attribute that mutual fund investors value, investors should have a tendency to repeat purchases from the same fund family. Considering this tendency from an aggregate basis, I find that about 27% of the purchases of mutual funds in the sample are purchases of funds that belong to a fund family in which the household has previously invested. Examining this tendency on an annual basis, the percentage is consistent throughout the sample period, never dropping below 24.3% in any sample year.

To more thoroughly test whether investors choose mutual funds in the same mutual fund family for repeat purchases, I conduct a probit regression each month and then aggregate across the months using a Fama-MacBeth (1973) analysis. The independent variable in these regressions is whether a particular mutual fund family was selected by an individual investor initiating a new position in a mutual fund. I treat each purchase as an independent event, and I consider only initiations made by households who have previously held at least one other fund in their brokerage account. For every new initiation, the dependent variable takes a value of 1 for the family of the selected fund and a value of zero for all other families in the sample. Because I do not have a list of which mutual fund families were available to the discount brokerage firm customers, I define the sample of available fund families as all families held by at least one individual investor in the year of the purchase.

The variable of interest is the independent variable *previous family ownership*, which takes a value of one if the investor has previously held a fund from that mutual fund family and zero otherwise. I also control for other mutual fund attributes that have been found to be determinants of investment choices in mutual funds or mutual fund families. (See, for example, Sirri and Tufano (1998), Barber, Odean and Zheng (2005), and Gallaher, Kaniel and Starks (2006)). These variables are defined as follows: Family total expenses are estimated as the weighted average of the total expenses for each fund in the family, where each fund is weighted by its TNA. Following Sirri and Tufano (1998), total expenses for a particular fund is defined as expense ratio + 1/7 * front end load fee (if any). The "1/7" term reflects an average holding period of 7 years for equity mutual funds, as reported by the Investment Company Institute (1991). To address concerns that the average holding period in our unique sample may differ, we replace total expenses with both expense ratio and front end load included separately as control variables in certain specifications. Including both controls requires no assumptions on the average holding period, as the two variables encompass any linear combination that could be included as a control variable.

I calculate an objective-adjusted return for each family as the weighted average 12-month objective-adjusted return for each fund in the family, where funds are weighted by their TNA. As defined previously, the objective-adjusted return for a fund is the return minus the equal-weighted average 12-month buy and hold return for all funds in the same objective class.[44] I also include two imperfect proxies for family reputation, family size, calculated as the family's TNA (the sum of the TNA of all sample funds belonging to the family), and family age, calculated according to the age of the oldest fund in the family. As a firm's reputation grows, its product sells more (Milgrom and Roberts, 1986), which in the mutual fund industry implies that the size of the assets under management should grow. As alternative controls for the fact that some families offer more fund choices than do others, thus increasing the probability that their funds would be selected, I include the number of funds offered by the family and the number of unique ICDI objectives offered by the funds in the family.[45]

Estimated marginal coefficients and t-statistics from this analysis are provided in Table 4.5.[46] In the first model of Table 4.5, I include only the variable that indicates whether the investor already holds a fund in the same family. As conjectured, the coefficient of 0.072 is positive and both economically and statistically significant. This coefficient indicates that a fund family's probability of being selected by a particular investor increases by 7.2% if the investor has previously held a fund from that family. To put this number in perspective, the average unconditional probability that a family is selected is only 0.90%, and the constant from the regression, representing the probability that a family is selected when all control variables are held at their means, is 0.7%. Thus, a family is more than 10 times more likely to be selected if the investor has prior

[44] For robustness, we also control for each family's raw 12-month return.
[45] The number of funds and number of objectives variables also partially address the fact that firms with more offerings are more likely to produce star funds, which has been shown by Nanda, Wang and Zheng (2004), Massa (2001) and others to increase fund flow to the entire family.
[46] All t-statistics are corrected for serial correlation up to 12 lags (one year) using the Newey-West correction.

experience with it. After including the control variables in the remaining models of the table, I find that the coefficient on the same family indicator variable remains positive and strongly significant in every specification. In fact, the coefficient estimate on the previous family ownership variable is at least six times larger than the estimated constant in every specification of Table 4.5.

To facilitate accurate comparisons between my variable of interest, previous family ownership, and other variables in the probit regressions, Table 4.6 reports standardized coefficients for each of the five regressions from Table 4.5. Since each coefficient in Table 4.5 represents the time-series average marginal probability associated with a one unit change in the independent variable, estimates are standardized by multiplying the coefficient by the time-series average of its cross-sectional standard deviation.[47] Thus, each coefficient in Table 4.6 represents the marginal change in probability associated with a one standard deviation change in the independent variable. The results in Table 4.6 underscore the economic significance of the same family indicator variable. Indeed, a one standard deviation increase in the same family indicator increases a family's probability of being chosen by between 0.584% and 1.41%, depending on the specification. By comparison, a one standard deviation increase in family objective-adjusted return or TNA increases the family's probability of being chosen by less than 0.3%. In fact, the effect of a one standard deviation change in previous family ownership is larger than the standardized effect of any other variable in all specifications.

[47] For ease of readability, the standardized coefficients in Table 4.6 are also converted to percentages by multiplying by 100.

4.4.2 Controlling for Investor's Experienced Return

The result that individuals are much more likely to purchase funds from families with which they have previous experience is consistent with the hypothesis that family reputation is an important component of mutual fund choice. However, since I define reputation as a belief in the competence of a fund manager or family of managers that arises separately from the effects of recent past performance, it is important to adequately control for the effect of near past performance in the regression analysis. Fortunately, the nature of the discount brokerage dataset allows us to test whether the significant positive effect of previous family ownership on future purchases exists even for customers whose experience with a family is not particularly good. If investors continue to be repeat purchasers even for those families with which they experience negative abnormal returns, it is a strong indication that some belief that is independent of performance, i.e. reputation, is driving their behavior.

To conduct such a test, I define two new variables, objective-adjusted return experienced and raw return experienced, for households that have previously held at least one other fund in their brokerage accounts. These new variables measure the objective-adjusted and raw past buy and hold returns that an investor experiences with a family as of the end of the month prior to the current mutual fund purchase.[48, 49] For families never held by the investor prior to the current purchase, both return experience variables take

[48] If the investor initiates a position in a fund during the middle of a month, the entire monthly return for that fund is attributed to the investor's experience. This practice is necessary because although we can identify the exact date that an investor initiates a mutual fund position, data on mutual fund returns is only available from CRSP at the monthly frequency.

[49] When multiple funds from the same family are held in a month, I calculate the households' weighted average return with the family by weighting each fund by the market value of the investors' equity in the fund.

values of zero. Since these return experience variables only take non-zero values when the investor has previously invested with a family, they both act as interaction variables between the previous family ownership indicator and the experienced return.

Including the two return experience measures with the other independent variables, I repeat the monthly Fama-Macbeth probit methodology from Section 4.4.1. The standardized coefficients and Newey-West corrected t-statistics from these regressions are presented in Table 4.7. Not surprisingly, columns one and two of Table 4.7 reveal that the objective-adjusted return and raw return, respectively, experienced by an investor with a particular family have a positive and statistically significant impact on the probability that the investor purchases another fund from the family. It is noteworthy, however, that the magnitude of the return experience variables is small when compared to the size of the previous family ownership indicator in Table 4.6. For instance, the standardized coefficient on the previous ownership variable in column five of Table 4.6 (0.584) is roughly ten times the magnitude of the objective-adjusted return experienced variable in column one of Table 4.7 (0.058).

To further demonstrate this disparity, I include the return experience variables and the previous family ownership indicator simultaneously in columns three and four of Table 4.7. The results clearly demonstrate that the impact of the previous ownership indicator is not being driven solely by the investor's previous return experience with the family. In column three, a one standard deviation change in the previous ownership variable changes the probability that a family is chosen by 0.571%, while a one standard deviation change in objective-adjusted return experienced affects the probability by only

0.015%. Similarly, in column four, a one standard deviation change in the previous ownership variable changes the probability that a family is chosen by 0.521%, while a one standard deviation change in raw return experienced affects the probability by only 0.019%.

4.4.3 Investor Demographics

The work of Barber and Odean (2001), among others, has shown that demographics, such as the investor's gender, can have a large impact on the trading behavior of individuals. In this section, I investigate the role that investor demographics play in affecting the relationship between fund family membership and the revealed choices of individual mutual fund investors.

Figures 4.1 through 4.3 provide a first glimpse at the differences among demographic groups, reporting the percentage of same family purchases for several different shareholder characteristics. Figure 4.1 is divided by investor age group, where the age of the head of household is measured in two year increments.[50] Figure 4.2 is divided by the gender of the head of household, and Figure 4.3 is divided by household income levels.[51] In all three figures, a same family purchase is defined as an initiation in a mutual fund that belongs to a family with which the investor has previously held another fund. A new position is defined as a possible same family purchase if the transaction is made by an investor who has previously held any mutual fund in her brokerage account.

---

[50] Thus a 22-year-old and a 23-year-old would each be assigned an age of 22 in the dataset, while a 24-year-old or 25-year-old would be assigned an age of 24.

[51] Investors in the database are grouped by income level into one of nine groups. Income is thus reported as a number between one and nine, with one corresponding to the lowest income and nine corresponding to the highest.

The same family purchase percentage is then defined within each demographic group as total same family purchases divided by total possible same family purchases. From Figure 4.1, it appears that a larger proportion of possible same family purchases by older investors are same family purchases. In contrast, Figures 4.2 and 4.3 show that little economic difference apparently exists between males and females or among different income levels in their tendency to purchase into previously-held fund families.

While the figures provide a quick glance at the data on a univariate basis, these figures should be interpreted with caution. A key limitation of the same family purchase percentage calculation is that it involves no control for the number of funds held previously by a particular investor. An investor who has previous experience with a large number of funds from many different families will mechanically be more likely to initiate new positions in funds from families that he has previously held. To control for this portfolio size effect, as well as to better isolate the importance of previous family ownership from other considerations such as family size, I examine the interaction between demographic variables and the previous family ownership indicator using the Fama-Macbeth regression framework described in Section 4.4.1.

The results from interacting investor demographics with the previous family ownership variable are provided in Table 4.8.[52] Columns one and three of Table 4.8 show

---

[52] Note that it is not necessary to include the demographic variables separately (without interaction) in the regression. Recall that for each qualifying initiation, each family available to sample investors in that year receives a one or zero depending on whether it was the family selected. While family-level variables display variation within each initiation, there is no variation in the investor information for each initiation. Thus, the probability assigned to each investor characteristic, in the absence of a constant, would be 1/n, where n is the number of families available to the investor. Since the regressions already include a constant, the coefficient estimate on any investor-level variable included separately in the regression will be approximately zero.

that there is no significant difference in the impact that previous family ownership has on the probability of choosing a family for investors of different gender or income level. Column two, however, does reveal a significant difference in the effect of previous family ownership for investors of different age. Interestingly, the results reveal that the same family ownership effect is significantly weaker for older investors. This finding reveals that the pattern suggested by Figure 4.2 does not remain after controlling for other investor and family characteristics in regressions.[53]

As a final test, I proxy for investment experience by measuring the length of time between the current mutual fund initiation and the date that the household first opens its account at the discount brokerage. The results from interacting the investor's tenure with the brokerage with previous family ownership in the regression are presented in column four of Table 4.8. The results reveal that previous family ownership plays a significantly larger role for more experienced investors. These results, which go opposite of the findings for the age interaction, indicate that the brokerage tenure and age variables capture two different aspects of investor behavior.

4.5 ALTERNATIVE EXPLANATIONS

While the portfolio concentration and repeat buyer results presented in previous sections are consistent with a reputation story, these results are also consistent with alternative explanations. One such alternative explanation is that investors repeatedly buy

---

[53] A possible cause for the apparently conflicting findings related to investor age in Figure 4.2 and Table 4.8 is that the same family purchase percentage ignores the number of families with which the investor has previous experience. A Pearson correlation test reveals that head of household age and the number of families previously held has a positive correlation of 0.062 (p-value < 0.0001).

from the same families not because reputation has cemented their beliefs in the competence of the fund managers but because the family's overall investment objective is aligned with the preferred style of the investor. For instance, an investor who prefers investing in value stocks may concentrate her investments in a mutual fund family that specializes in equity value funds.

Another alternative explanation for the repeat buying results is familiarity. Familiarity with a mutual fund family may result from the family's reputation through such channels as advertising, media coverage or word of mouth, and in this sense familiarity and reputation are similar concepts. However, there is one specific and direct channel by which investors can become familiar with the funds of a particular family with which they have prior experience that could be driving the repeat buyer results--shared prospectuses. It is common practice for fund families to include member funds belonging to a particular investment objective in the same prospectus. Thus, investors who hold a fund from a particular family will almost certainly be somewhat more familiar with the family's other offerings within the same objective class. If the repeat buying results derive mainly from instances in which investors purchase funds of the same objective from the same family, then whether my results reveal information about the true nature of reputation on fund choice is questionable.

To investigate the validity of both alternative explanations, style matching and familiarity resulting from shared prospectuses, I repeat the regressions of Section 4.4.1 on the subset of initiations in which the investor purchases a fund from an objective class that she has never previously held. If the importance of previous family ownership is not diminished by examining only this subset of purchases, then it is unlikely that the previous results are simply an artifact of style alignment between investor and family or familiarity resulting from shared prospectuses.

Table 4.9 reports standardized coefficient estimates and Newey-West corrected t-statistics for the subsample of initiations in new objective categories.[54] Comparing Table 4.9 to Table 4.6 reveals that the effect of previous family ownership is not diminished (and in fact is slightly larger) for the new objective subsample. In univariate regressions, the coefficient estimate for a one standard deviation increase in previous family ownership is 1.56% for investors expanding to new objectives, compared to 1.41% for the full sample. After including family-level controls, the coefficient estimate remains a positive and highly statistically significant 0.637%, as compared to 0.584% in Table 4.6. These results indicate that the importance of previous family ownership on individuals' mutual fund decisions does not derive from a preference for families that specialize in particular objectives or from familiarity related to the sharing of prospectuses by family funds belonging to the same objective category.

4.6 FAMLY REPUTATION AND FLOW-PERFORMANCE SENSITIVITY

As argued by Ippolioto (1992) and others, investors rely on past performance as a signal of mutual fund quality. Hence, the sensitivity of flow to past returns is a measure of how investors update their beliefs about funds. Mailath and Samuelson (2001) show that a strong reputation (i.e. when consumers have strong posterior expectations that the

[54] Fund objective is determined using the ICDI objective code from the CRSP mutual fund database. The ICDI objective code contains 23 total objective classifications.

firm is competent) will result in consumers' beliefs changing only gradually when faced with additional noisy signals of competence. In the context of the current study, this gives rise to the hypothesis that investors will be slower to update their beliefs about mutual funds belonging to families with more established reputations.

To test this hypothesis, I examine whether the response of fund flows of individual investors to past returns is significantly different for funds belonging to families with stronger reputations. Specifically, I conduct a panel regression with robust standard errors of aggregate individual order imbalances, scaled by size, on lagged mutual fund characteristics. I estimate the following model:

$$(buymv\text{-}sellmv)_{i,t}/TNA_{i,t-1} = \alpha + \beta(fund\ controls_{i,t-1}) + \gamma(fam\ controls_{j,t-1}) + \delta(oaret12_{i,t-1}\ x\ reputation\ proxies_{j,t-1}) + \varepsilon, \quad (4)$$

where $buymv_{i,t}$ is the market value of purchases by individuals of fund i in year t and $sellmv_{i,t}$ is the market value of sells by individuals of fund i in year t. Because several fund-level variables have been shown to be important determinants of fund flow, I estimate flow at the individual fund level, rather than for the entire family. I then include both fund- and family-level variables as controls in the regression. The fund-level variables are fund size (TNA), fund expenses (expenses + 1/7 * front end load fee, measured in basis points), turnover ratio, fund age, and objective-adjusted return (calculated as the buy and hold return for the previous 12-month calendar year minus the equal-weighted average 12-month buy and hold return for all funds in the same objective class). The family-level variables are family size (family TNA), family age (defined as the maximum age among the funds comprising the family), family objectives (the number of unique ICDI objectives offered by the family), family-level total expenses, and objective-adjusted return. The latter two family variables are defined as weighted averages of the corresponding variables for member funds in each year, where member funds are weighted within each family by their TNA. Finally, I include interactions of objective-adjusted return with two proxies for family reputation, family size and family age.

Table 4.10 reports the coefficient estimates and robust t-statistics from the regressions. To begin, note that the estimate on fund objective-adjusted return is positive and significant in all regression specifications. This positive relationship between past return and flow confirms that mutual fund flows from sample individuals follow a similar pattern to aggregate flows examined previously in the literature (e.g., Sirri and Tufano 1998).

Of more importance to my analysis is the coefficient δ from equation (4) above. As stated in the introduction, we hypothesize that investors' beliefs about funds belonging to families with more established reputations will change more slowly than for funds lacking reputation. Thus, I predict that δ will be negative, as investors will be slower to update their beliefs about mutual funds belonging to families with established reputations.

The results of Table 4.10 confirm this intuition. The coefficients on the interactions of past objective-adjusted return with family size and family age are negative and significant at the five and ten percent significance levels, respectively. These findings indicate that individual investors' flows are significantly less sensitive to past returns for funds belonging to larger, older fund families, even after controlling for a host of fund- and family-level characteristics.

Before concluding, it should be noted that although I provide evidence that suggests reputation decays slowly, there is also anecdotal evidence that a scandal can cause a big loss in reputational capital and a very fast subsequent reaction. That is, in contrast to the growth due to a building of reputational capital, if a firm's reputation is seriously harmed, the size of the assets under management can be rapidly reduced. This was the case with investment management companies that were charged with allowing late trading in the mutual fund scandals of 2003. For example, Putnam lost 12% of its assets under management during November 2003 after being charged by federal and state regulators for civil fraud on October 28, 2003 (Wall Street Journal, 2003). As Jack Brennan, CEO of Vanguard Group, has been quoted as saying, "Client trust is the most important attribute to attaining success, even a small dose of doubt will overrun great marketing."

4.7 CONCLUSIONS

In this chapter, I examine the influence of fund family membership on the mutual fund selection decision. Analyzing individuals' mutual fund holdings and trades at a large discount brokerage firm, I provide evidence that is consistent with the hypothesis that mutual fund family reputation is an important factor in individual investor decisions. First, individuals tend to cluster their investments within particular families, even though they are purchasing funds through a fund supermarket for which there are no institutional advantages to remaining within a fund family. Second, the sample investors are significantly more likely to purchase funds from families in which they have already held funds and, thus, have previous experience. Moreover, this effect does not change substantially depending on the return that the investor previously experiences with the family, indicating that the repeat buyer effect is measuring something that is independent of past performance. Third, consistent with my reputation explanation, but inconsistent with alternative explanations such as family style preferences or familiarity arising from shared prospectuses within family objective classes, I find that the positive effect of previous family ownership on future fund selection obtains even for investors choosing funds with investment objectives that are new to the investor. Finally, consistent with the Mailath and Samuelson (2001) result that reputational beliefs dissipate slowly, I find that individuals' beliefs about funds belonging to older and larger families change slowly, as evidenced by decreased flow-performance sensitivity for these funds.

Table 4.1: Mutual Fund Holdings of Sample Households

This table contains summary statistics for the mutual fund holdings of sample households at a large discount brokerage for each year from 1991 to 1996. Sample statistics represent the equal-weighted average across households' mutual fund holdings as of January of each year. Panel A contains sample statistics regarding all households in the sample. Panel B contains summary statistics for those households that held mutual funds in their brokerage accounts. Panel C gives the distribution of mutual fund market value for households that held mutual funds in their brokerage accounts.

Panel A: All households

	Number of households	Number of households holding mutual funds	% of households holding mutual funds	Average portfolio % of mutual fund holdings
1991	56,023	8,798	15.7%	8.9%
1992	66,571	13,423	20.2%	10.4%
1993	56,999	14,272	25.0%	12.3%
1994	41,562	12,880	31.0%	14.5%
1995	33,478	10,953	32.7%	15.1%
1996	28,266	9,886	35.0%	16.3%

Panel B: Households holding mutual funds

	House -holds	Number of mutual funds per household		Number of mutual fund shares per household		Market value of mutual funds per household		Average portfolio % of mutual fund holdings	
		Mean	Max	Mean	Median	Mean	Median	Mean	Median
1991	8,798	2.15	26	2,149	434	17,969	6,905	56.8%	55.0%
1992	13,423	2.30	41	2,395	507	24,608	8,890	51.5%	47.2%
1993	14,272	2.62	99	3,064	619	29,922	10,392	48.9%	43.8%
1994	12,880	2.96	219	4,228	721	37,351	12,354	46.6%	41.0%
1995	10,953	3.03	356	8,806	765	36,976	11,610	45.7%	39.6%
1996	9,886	3.20	443	12,524	869	49,533	16,011	46.2%	40.3%

Panel C: Distribution of average mutual fund market value for participating households

	P5	P10	P25	P50	P75	P90	P95
1991	452	1,338	3,198	6,905	15,695	38,419	66,832
1992	984	1,679	3,915	8,890	20,629	52,086	98,140
1993	1,000	1,807	4,363	10,392	25,170	64,043	114,444
1994	1,132	2,171	5,206	12,354	30,822	77,880	144,489
1995	1,029	2,022	4,900	11,610	30,281	76,807	143,929
1996	1,386	2,699	6,455	16,011	41,649	108,234	194,031

Table 4.2: Mutual Fund Summary Statistics

Table 4.2 contains summary statistics from the CRSP mutual fund database as of the end of each calendar year from 1990 to 1996. Average Total Net Assets (TNA) is reported in millions of dollars. Age represents the average number of years since the funds first appeared in the CRSP database. Total fees is defined as expenses + 1/7 * front end load fee. Return is the buy and hold return for the 12-month calendar year. The objective-adjusted return (oa-return) is the return minus the equal-weighted average 12-month buy and hold return for all funds in the same objective class. Panel A contains statistics for all mutual funds in the CRSP database. Panel B contains statistics for funds that were held by at least one household during the sample.

Panel A: All CRSP-listed mutual funds

	N	TNA	Age	Total Fees	Return	OA-Return
1990	3,684	306.8	7.97	1.03%	0.92%	0.03%
1991	4,065	342.9	7.90	0.73%	18.26%	0.01%
1992	5,288	326.2	7.16	1.23%	6.15%	-0.11%
1993	6,913	322.1	6.29	1.23%	11.65%	-0.06%
1994	8,750	241.6	5.62	1.30%	-2.25%	-0.11%
1995	9,350	298.3	6.17	1.38%	17.16%	-0.07%
1996	10,125	332.0	6.55	1.42%	9.91%	-0.06%

Panel B: Mutual funds held by sample individuals

	N	TNA	Age	Total Fees	Return	OA-Return
1990	533	554.6	13.11	1.26%	-1.81%	0.60%
1991	567	721.4	13.32	1.06%	26.80%	1.16%
1992	631	839.7	13.13	1.25%	8.92%	1.78%
1993	683	1033.7	13.18	1.25%	17.01%	1.34%
1994	728	1000.5	13.58	1.26%	-1.78%	1.04%
1995	751	1308.7	14.18	1.27%	23.28%	1.19%
1996	770	1581.7	14.89	1.26%	14.03%	0.34%

Table 4.3: Fund Family Summary Statistics

Table 4.3 contains family-level summary statistics from the CRSP mutual fund database as of the end of each calendar year from 1990 to 1996. Family TNA is defined as the sum of the TNA of each member fund belonging to the management company in each year. TNA is reported in millions of dollars. Family age is defined as the maximum age among the funds comprising the family, rounded to the nearest year. Family-level expenses, NAV, total fees, turnover, return and objective-adjusted return are defined as weighted averages of the corresponding variables for member funds in each year, where member funds are weighted within each family by their TNA. Total fees is defined as expenses + 1/7 * front end load fee. Return is the buy and hold return for the 12-month calendar year. OA-Return is the return minus the equal-weighted average 12-month buy and hold return for all funds in the same objective class. Panel A contains statistics for all mutual fund families in the CRSP database. Panel B contains family statistics only for families that were held by at least one household during the sample.

Panel A: All fund families

	N	TNA	Age	Funds	Objectives	Total Fees	Return	OA-Return
1990	378	2,730	17.4	8.4	4.8	1.18%	-0.20%	-0.30%
1991	400	3,288	17.5	9.3	5.0	0.80%	21.28%	-1.12%
1992	435	3,496	16.4	11.6	5.4	1.37%	6.61%	-0.06%
1993	442	4,449	16.6	15.3	5.8	1.38%	11.47%	-0.44%
1994	473	4,271	17.3	18.3	5.8	1.41%	-1.49%	0.20%
1995	452	5,712	18.3	20.3	5.9	1.46%	19.92%	-1.11%
1996	457	6,790	18.4	21.5	5.9	1.46%	13.12%	-0.25%

Panel B: Fund families held by sample individuals

	N	TNA	Age	Funds	Objectives	Total Fees	Return	OA-Return
1990	143	5,170	23.1	12.9	6.3	1.10%	-0.72%	-0.09%
1991	148	6,340	23.3	14.1	6.6	0.86%	23.89%	0.29%
1992	150	7,422	23.4	18.2	7.2	1.26%	8.39%	1.61%
1993	152	9,182	23.5	24.9	7.6	1.26%	14.94%	0.88%
1994	160	9,091	24.2	30.4	7.6	1.29%	-1.52%	0.50%
1995	157	12,040	24.7	33.7	7.7	1.30%	21.70%	-0.33%
1996	158	14,558	25.1	34.9	7.6	1.32%	14.95%	0.72%

Table 4.4: Within-Family Clustering of Individuals' Mutual Fund Investments

Table 4.4 shows the concentration of sample individuals' mutual fund investments within mutual fund families. Sample households are divided by the number of different mutual funds that they hold at any time during the sample period. A column labeled "n funds in same family" reports the percentage of households within each group that hold at least n funds that are in the same mutual fund family. The final column reports the percentage of households within each group that concentrate all of their mutual fund holdings within the same family.

Mutual Funds Held	Households	Two Funds in Same Family	Three Funds in Same Family	Four Funds in Same Family	All Funds in Same Family
1	10,002	.	.	.	100.0%
2	5,064	32.3%	.	.	32.3%
3	3,188	50.6%	13.3%	.	13.3%
4	2,045	66.0%	21.8%	5.9%	5.9%
5	1,404	77.5%	30.6%	10.4%	3.6%
6	1,050	84.1%	38.0%	15.4%	2.0%
7	747	88.4%	44.8%	19.9%	0.9%
8	606	93.4%	52.0%	26.7%	1.3%
9	462	95.5%	58.0%	28.1%	1.5%
10	366	96.7%	64.2%	36.3%	2.5%
11+	1,937	99.3%	84.5%	60.4%	0.2%

Table 4.5: Probability of Choosing a Fund Family

This table reports Fama-Macbeth estimates of the marginal probabilities from monthly probit regressions where the independent variable is whether a particular mutual fund family was selected. For every new initiation by a household that has previously held a mutual fund, the dependent variable takes a value of 1 for the family of the selected fund and a value of zero for all other families in the sample, which consists of all fund families held by at least one individual investor in the year of the initiation. The independent variable *previous family ownership* takes a value of one if the investor has previously held a fund from that mutual fund family. Total expenses for a particular fund is defined as expense ratio + 1/7 * front end load fee. A fund's objective adjusted return is the 12-month buy and hold return minus the equal-weighted average return for all funds in the same objective class. Family total expenses, oa-ret12, front load, expense ratio, and ret12 are the weighted average of the total expenses, objective-adjusted return, front load fee, expense ratio and raw return for each fund in the family, where each fund is weighted by its TNA. Family TNA is the sum of the TNA of all sample funds belonging to the family. Family objectives is the number of unique ICDI objectives offered by the funds in the family. Family age is the age of the family, and family funds is the number of funds belonging to the family. Fama-Macbeth t-stats are corrected for serial correlation up to 12 lags using Newey-West and are reported in italics.

Probit Regression: Fund Family Purchase = 1
Avg. Unconditional Probability = 0.009025

	(1)	(2)	(3)	(4)	(5)
Intercept	0.007	0.005	0.005	0.005	0.005
	10.73	*11.60*	*11.56*	*12.36*	*12.82*
Previous family ownership	0.072	0.036	0.039	0.043	0.030
	10.51	*13.01*	*11.76*	*17.92*	*15.04*
Family total expenses		-0.534	-0.585	-0.594	
		-11.30	*-8.14*	*-11.94*	
Family oa-ret12		0.029	0.037	0.032	
		4.47	*4.62*	*4.64*	
Family TNA		7.8E-08			7.8E-08
		5.93			*5.35*
Family age		-3.0E-05		7.2E-06	7.6E-06
		-1.56		*0.64*	*0.50*
Family objectives			3.7E-04		
			9.83		
Family funds				3.7E-05	
				3.69	
Family front load					-0.210
					-6.52
Family expense ratio					-0.141
					-3.99
Family ret12					0.013
					2.41
Avg. Pseudo-R^2	0.087	0.140	0.131	0.129	0.150

Table 4.6: Probability of Choosing a Fund Family

This table reports standardized Fama-Macbeth estimates of the marginal probabilities from monthly probit regressions where the independent variable is whether a particular mutual fund family was selected. Each coefficient estimate is standardized by multiplying the time-series average marginal probability by its average cross-sectional standard deviation. For every new initiation by a household that has previously held a mutual fund, the dependent variable takes a value of 1 for the family of the selected fund and a value of zero for all other families in the sample, which consists of all fund families held by at least one individual investor in the year of the initiation. The independent variable *previous family ownership* takes a value of one if the investor has previously held a fund from that mutual fund family. Total expenses for a particular fund is defined as expense ratio + 1/7 * front end load fee. A fund's objective-adjusted return is the 12-month buy and hold return minus the equal-weighted average return for all funds in the same objective class. Family total expenses, oa-ret12, front load, expense ratio, and ret12 are the weighted average of the total expenses, objective adjusted return, front load fee, expense ratio and raw return for each fund in the family, where each fund is weighted by its TNA. Family TNA is the sum of the TNA of all sample funds belonging to the family. Family objectives is the number of unique ICDI objectives offered by the funds in the family. Family age is the age of the family, and family funds is the number of funds belonging to the family. Fama-Macbeth t-stats are corrected for serial correlation up to 12 lags using Newey-West and are reported in italics. All coefficient estimates are converted to percentages.

Probit Regression: Fund Family Purchase = 1 Avg. Unconditional Probability = 0.9025%					
	(1)	(2)	(3)	(4)	(5)
Intercept	0.709	0.521	0.512	0.534	0.450
	10.73	*11.60*	*11.56*	*12.36*	*12.82*
Previous family ownership	1.413	0.716	0.775	0.851	0.584
	10.51	*13.01*	*11.76*	*17.92*	*15.04*
Family total expenses		-0.289	-0.317	-0.322	
		-11.30	*-8.14*	*-11.94*	
Family oa-ret12		0.219	0.272	0.237	
		4.47	*4.62*	*4.64*	
Family TNA		0.212			0.214
		5.93			*5.35*
Family age		-0.066		0.016	0.017
		-1.56		*0.64*	*0.50*
Family objectives			0.231		
			9.83		
Family funds				0.171	
				3.69	
Family front load					-0.410
					-6.52
Family expense ratio					-0.063
					-3.99
Family ret12					0.126
					2.41
Avg. Pseudo-R^2	0.087	0.140	0.131	0.129	0.150

Table 4.7: Effect of Experienced Return on Choosing a Fund Family

This table reports standardized Fama-Macbeth estimates of the marginal probabilities from monthly probit regressions where the independent variable is whether a particular mutual fund family was selected. Each coefficient estimate is standardized by multiplying the time-series average marginal probability by its average cross-sectional standard deviation. For every new initiation by a household that has previously held a mutual fund, the dependent variable takes a value of 1 for the family of the selected fund and a value of zero for all other families in the sample, which consists of all fund families held by at least one individual investor in the year of the initiation. The independent variable *previous family ownership* takes a value of one if the investor has previously held a fund from that mutual fund family. A fund's objective-adjusted return is the 12-month buy and hold return minus the equal-weighted average return for all funds in the same objective class. Family oa-ret12, front load, and expense ratio are the weighted average of the total objective-adjusted return, front load fee, and expense ratio for each fund in the family, where each fund is weighted by its TNA. *OA-return experienced* and *raw return experienced* measure the objective-adjusted and raw past buy and hold returns that the investor experienced with each family as of the end of the month prior to the current mutual fund purchase. For families never held by the investor prior to the current purchase, both return experience variables take values of zero. Family TNA is the sum of the TNA of all sample funds belonging to the family. Fama-Macbeth t-stats are corrected for serial correlation up to 12 lags using Newey-West and are reported in italics. All coefficient estimates are converted to percentages.

Probit Regression: Fund Family Purchase = 1 Avg. Unconditional Probability = 0.9025%				
	(1)	(2)	(3)	(4)
Intercept	0.488	0.481	0.446	0.447
	17.34	*16.75*	*14.25*	*14.29*
OA-Return experienced	0.058		0.015	
	5.80		*3.42*	
Raw return experienced		0.161		0.019
		3.49		*5.56*
Previous family ownership			0.571	0.521
			17.64	*15.94*
Family oa-ret12	0.193	0.186	0.169	0.169
	4.64	*4.73*	*4.66*	*4.70*
Family TNA	0.274	0.245	0.206	0.205
	6.31	*6.06*	*5.64*	*5.54*
Family age	0.038	0.027	0.010	0.010
	0.88	*0.65*	*0.29*	*0.28*
Family front load	-0.521	-0.482	-0.406	-0.405
	-7.67	*-7.49*	*-6.77*	*-6.75*
Family expense ratio	-0.087	-0.085	-0.066	-0.066
	-3.87	*-4.11*	*-3.69*	*-3.64*
Avg. Pseudo-R^2	0.107	0.124	0.151	0.151

Table 4.8: Demographics and the Probability of Choosing a Fund Family

This table reports standardized Fama-Macbeth estimates of the marginal probabilities from monthly probit regressions where the independent variable is whether a particular mutual fund family was selected. For every new initiation by a household that has previously held a mutual fund, the dependent variable takes a value of 1 for the family of the selected fund and a value of zero for all other families in the sample, which consists of all fund families held by at least one individual investor in the year of the initiation. The independent variable *previous family ownership* takes a value of one if the investor has previously held a fund from that mutual fund family. A fund's objective adjusted return is the 12-month buy and hold return minus the equal-weighted average return for all funds in the same objective class. Family oa-ret12, front load, and expense ratio are the weighted averages of the objective adjusted return, front load fee, and expense ratio for each fund in the family, where each fund is weighted by its TNA. Family TNA is the sum of the TNA of all sample funds belonging to the family. The demographic variables male, age, and income reflect self-reported demographic information for the head of household for each individual account. Brokerage tenure is measured as the number of days between the individual's account opening and the current fund initiation. Fama-Macbeth t-stats are corrected for serial correlation up to 12 lags using Newey-West and are reported in italics. Each coefficient estimate is standardized by multiplying the time-series average marginal probability by its average cross-sectional standard deviation and then converted to a percentage by multiplying by 100.

Probit Regression: Fund Family Purchase = 1				
	(1)	(2)	(3)	(4)
Avg. unconditional probability	*0.8982%*	*0.8977%*	*0.8981%*	*0.9025%*
Intercept	0.422	0.422	0.424	0.446
	17.73	*18.10*	*18.35*	*14.21*
Previous family ownership	0.740	1.238	0.776	0.498
	20.27	*18.05*	*11.26*	*9.61*
Family oa-ret12	0.159	0.157	0.158	0.170
	4.90	*4.93*	*4.93*	*4.72*
Family TNA	0.201	0.199	0.202	0.207
	5.63	*5.68*	*5.75*	*5.63*
Family age	0.010	0.010	0.010	0.011
	0.27	*0.26*	*0.27*	*0.30*
Family front load	-0.393	-0.393	-0.394	-0.405
	-6.84	*-6.81*	*-6.82*	*-6.77*
Family expense ratio	-0.051	-0.053	-0.053	-0.065
	-2.74	*-2.82*	*-2.79*	*-3.61*
Prev fam own * male	0.006			
	1.06			
Prev fam own * age		-0.063		
		-7.25		
Prev fam own * income			-0.007	
			-1.04	
Prev fam own * brokerage tenure				0.017
				3.05
Avg. Pseudo-R2	0.164	0.164	0.163	0.151

Table 4.9: Probability of Choosing a Fund Family for New Objectives

This table reports standardized Fama-Macbeth estimates of the marginal probabilities from monthly probit regressions where the independent variable is whether a particular mutual fund family was selected. For every new initiation by a household that has previously held a mutual fund but never held a fund from the same ICDI objective class as the current purchase, the dependent variable takes a value of 1 for the family of the selected fund and a value of zero for all other families in the sample, which consists of all fund families held by at least one individual investor in the year of the initiation. The independent variable *previous family ownership* takes a value of one if the investor has previously held a fund from that mutual fund family. A fund's objective adjusted return is the 12-month buy and hold return minus the equal-weighted average return for all funds in the same objective class. Family oa-ret12, front load, expense ratio, and ret12 are the weighted averages of the objective adjusted return, front load fee, expense ratio and raw return for each fund in the family, where each fund is weighted by its TNA. Family TNA is the sum of the TNA of all sample funds belonging to the family. Fama-Macbeth t-stats are corrected for serial correlation up to 12 lags using Newey-West and are reported in italics. Each coefficient estimate is standardized by multiplying the time-series average marginal probability by its average cross-sectional standard deviation and then converted to a percentage by multiplying by 100.

Probit Regression: Fund Family Purchase = 1 Avg. Unconditional Probability = 0.9038%			
	(1)	(2)	(3)
Intercept	0.709	0.431	0.432
	10.38	*14.38*	*12.74*
Previous family ownership	1.557	0.646	0.637
	19.42	*20.47*	*25.66*
Family oa-ret12		0.135	
		3.82	
Family ret12			0.095
			1.76
Family TNA		0.204	0.209
		5.40	*5.08*
Family age		0.003	0.014
		0.07	*0.37*
Family front load		-0.399	-0.404
		-5.95	*-5.85*
Family expense ratio		-0.105	-0.097
		-4.12	*-4.75*
Avg. Pseudo-R2	0.094	0.161	0.161

Table 4.10: Flow-performance sensitivity regressions

This table contains OLS regression results for the effect of lagged mutual fund characteristics on individual order imbalance scaled by size. I estimate the following model:

$$(buymv\text{-}sellmv)_{i,t}/TNA_{i,t-1} = \alpha + \beta(fund\ controls_{i,t-1}) + \gamma(fam\ controls_{j,t-1}) + \delta(oaret12_{i,t-1}\ x\ reputation\ proxies_{j,t-1}) + \varepsilon,$$

where $buymv_{i,t}$ is the market value of purchases by individuals of fund i year t and $sellmv_{i,t}$ is the market value of sells by individuals of fund i in year t. TNA is measured in millions of dollars. Total expenses is defined as expenses + 1/7 * front end load fee and is measured in basis points. Return is the buy and hold return for the 12-month calendar year. OA-Return is the return minus the equal-weighted average 12-month buy and hold return for all funds in the same objective class. Family TNA is defined as the sum of the TNA of each member fund belonging to the management company in each year. Family age is defined as the maximum number of years of existence for any of the funds comprising the family. Family-level total expenses, return and objective-adjusted return are defined as weighted averages of the corresponding variables for member funds in each year, where member funds are weighted within each family by their TNA. The model also includes interaction terms between the family objective-adjusted returns and family size, age, and objectives.

Dependent Variable: Value Buy Imbalance					
	(1)	(2)	(3)	(4)	(5)
Intercept	1,116	852	767	679	2,226
	3.15	*1.70*	*1.65*	*1.45*	*2.15*
Oaret12	12,093	14,667	12,836	21,234	35,631
	2.36	*2.30*	*2.69*	*2.62*	*2.11*
TNA		-0.11	-0.05	-0.05	-0.04
		-2.41	*-1.57*	*-1.50*	*-1.60*
Turnover		-290.27	-192.79	-191.92	-297.24
		-0.85	*-0.67*	*-0.67*	*-0.90*
Total fund expenses		11.36	6.24	6.51	7.72
		1.69	*1.09*	*1.13*	*1.30*
Fund age		-53.15	-62.16	-60.96	-32.82
		-2.02	*-2.01*	*-2.00*	*-1.87*
Family TNA			-0.01	-0.01	
			-2.17	*-1.90*	
Family age					-43.38
					-2.06
Family expenses			11.38	11.19	6.18
			1.00	*0.98*	*0.64*
Family oaret12			8,884	1,632	-2,802
			0.70	*0.14*	*-0.35*
Oaret12 x family TNA				-0.09	
				-2.35	
Oaret12 x family age					-492.73
					-1.73
R^2	0.002	0.005	0.006	0.007	0.009

Figure 4.1: Same Family Purchase Percentage by Investor Age

Figure 4.1 reports the percentage of same family purchases for eight different investor age groups. A *same family purchase* is a new position in a mutual fund that belongs to a family in which the investor has previously held another fund. A new position is a *possible same family purchase* if the transaction is made by an investor who has previously held any mutual fund in her brokerage account. The *same family purchase percentage* is defined within each age classification as total same family purchases divided by total possible same family purchases.

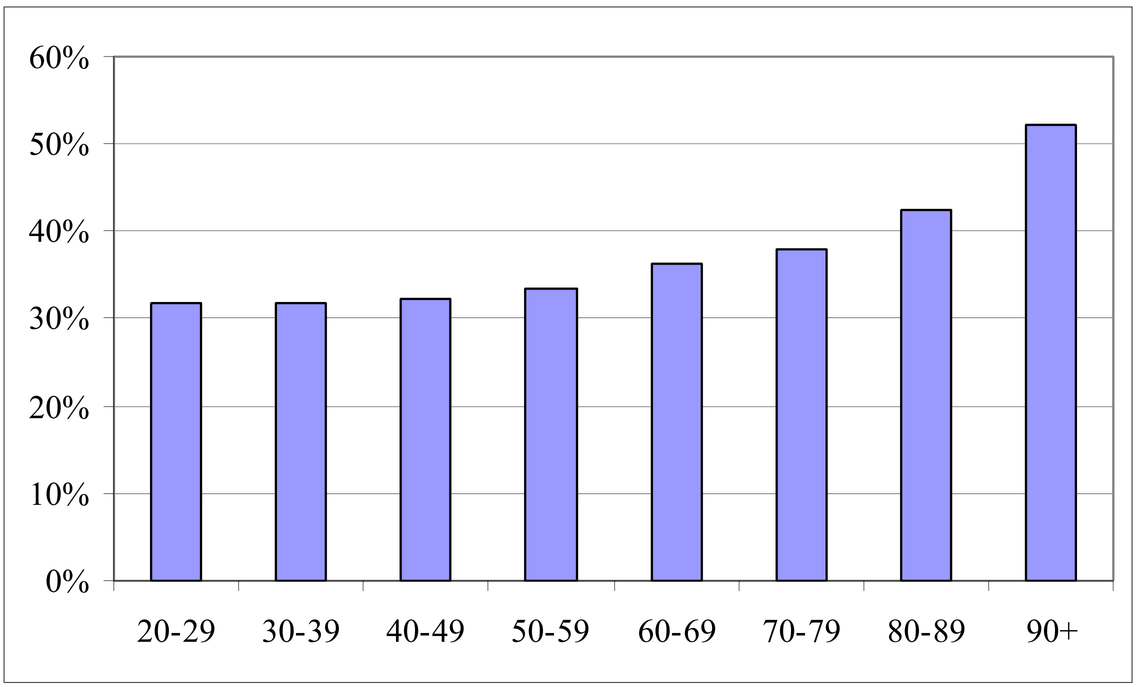


Figure 4.2: Same Family Purchase Percentage by Investor Gender

Figure 4.2 reports the percentage of same family purchases for females and males. A *same family purchase* is a new position in a mutual fund that belongs to a family in which the investor has previously held another fund. A new position is a *possible same family purchase* if the transaction is made by an investor who has previously held any mutual fund in his or her brokerage account. The *same family purchase percentage* is defined as total same family purchases divided by total possible same family purchases.

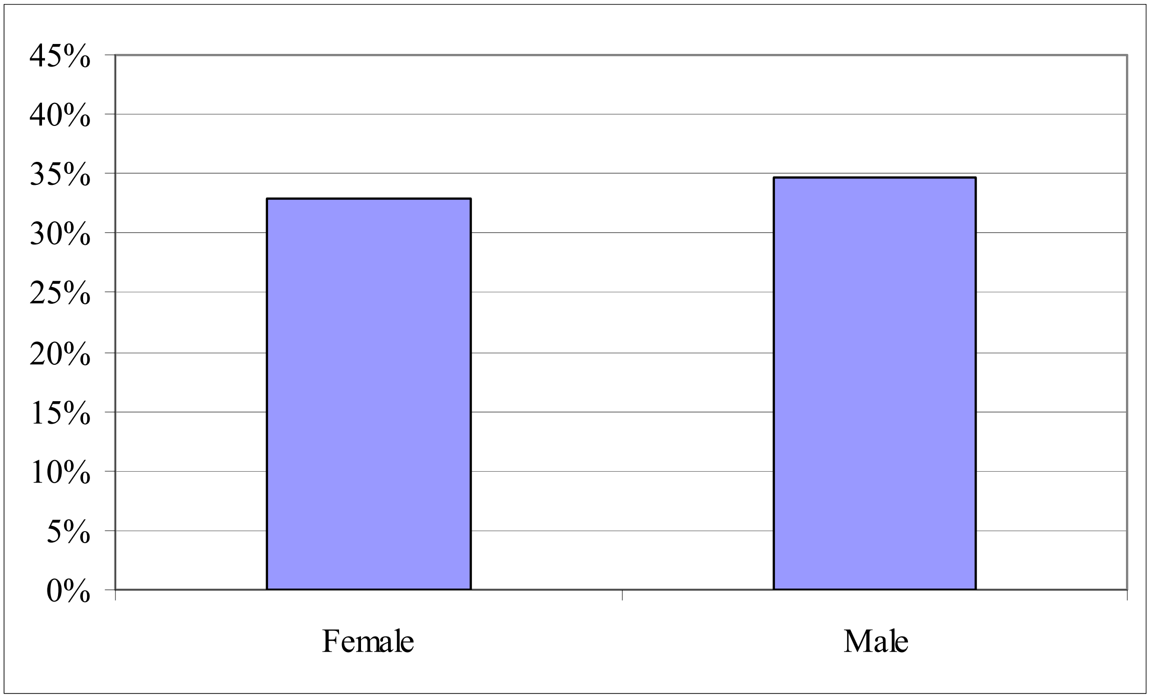
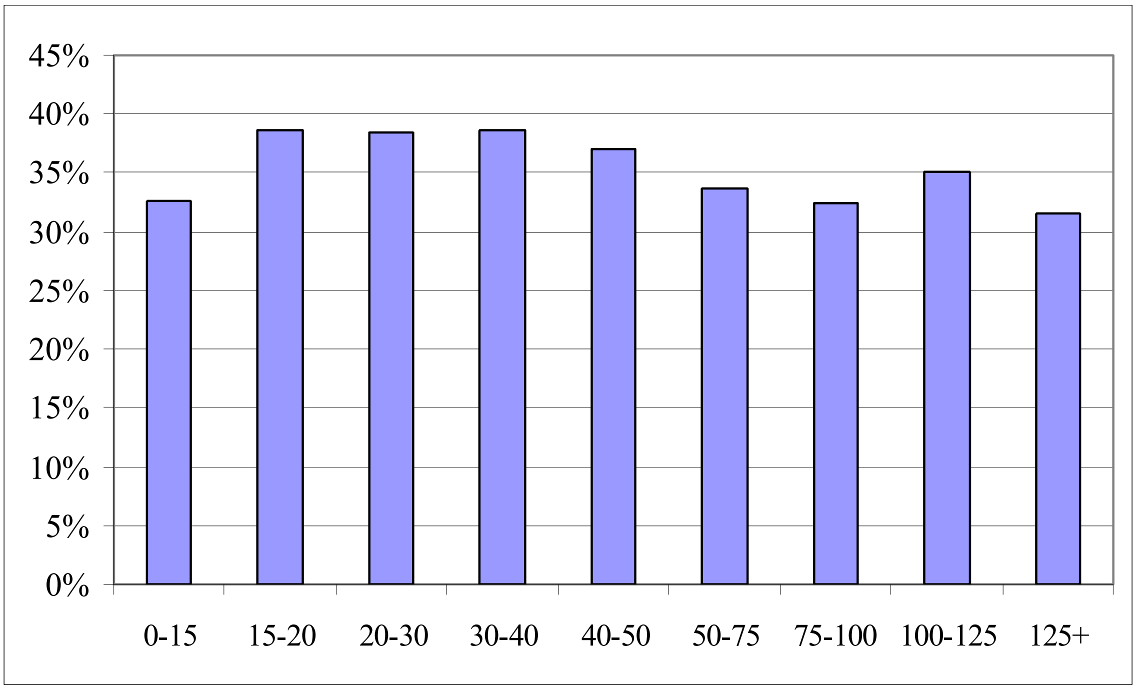


Figure 4.3: Same Family Purchase Percentage by Investor Income

Figure 4.3 reports the percentage of same family purchases for nine different investor income groups. Household income is reported in thousands of dollars. A *same family purchase* is a new position in a mutual fund that belongs to a family in which the investor has previously held another fund. A new position is a *possible same family purchase* if the transaction is made by an investor who has previously held any mutual fund in her brokerage account. The *same family purchase percentage* is defined within each income classification as total same family purchases divided by total possible same family purchases.



CHAPTER 5: CONCLUSION

This dissertation examines how individuals formulate their financial investment decisions and the importance of these decisions to the financial marketplace. Chapter 3 focuses on the importance of attention in influencing the common stock selections of individuals and shows that attention can have a significant impact on the returns of attention-grabbing equities. Consistent with other findings in the attention literature, I document that individuals' buy-sell imbalances for stocks that hit a 52-week high or low on the previous day are significantly more positive than for all other stocks.

In addition, I find that stocks experiencing 52-week highs and lows earn statistically and economically significant abnormal returns on days immediately following these events, and this effect remains positive and significant after controlling for various event-day, firm-level characteristics such as return, turnover, volatility and liquidity. I motivate these tests by arguing that 52-week highs and lows, in contrast to other attention-grabbing events that have been studied, represent relatively uninformative events. To solidify this claim, I repeat the return regressions after excluding from the sample a) the top 50% of all 52-week highs and lows as ranked on event-day absolute return, turnover or intraday volatility, and b) any 52-week highs or lows that occur within plus or minus three days of the underlying firm's earnings announcement. The positive and significant effect of 52-week highs and lows on future returns obtains even in this subsample of events which occur on relatively uneventful days, indicating that it is the occurrence of the event itself, and not correlation with an omitted information variable, that is driving my results.

Two additional results further demonstrate that the predictable abnormal returns following the event day are indeed caused by attention. First, I show that the return predictability is unique to 52-week highs and lows, as similarly-defined highs and lows relative to past prices which are not simultaneously 52-week events are shown to have no significant effect on next-day returns. Next, I document that on days when 52-week highs (lows) do the best job of narrowing an attention-driven trader's consideration set (i.e. when a relatively small number of 52-week highs (lows) occurs), the 52-week high (low) effect is significantly stronger. Again, this result is consistent with the hypothesis that the abnormal returns following 52-week high and low events are the result of attention-based trading.

Finally, I present evidence that the predictable returns following 52-week highs are confined to those events that exhibit positive individual buy-sell imbalances on the following day, suggesting that individual trading may be the conduit through which these short-term abnormal returns are generated.

In Chapter 4, I document the importance of mutual fund family affiliation for the mutual fund investment decisions of individual investors and argue that these results highlight the importance of individuals' beliefs about family reputation. Analyzing individuals' mutual fund holdings and trades at a large discount brokerage firm, I provide evidence that is consistent with the hypothesis that mutual fund family reputation is an important factor in individual investors' decisions. First, I show that individuals tend to cluster their investments within particular families, even though they are purchasing funds through a fund supermarket for which there are no institutional advantages to remaining within a fund family. Similarly, I find that sample investors are significantly more likely to purchase funds from families with which they have previous experience. Moreover, this effect does not change substantially depending on the return that the investor previously experiences with the family, indicating that the effect of previous family ownership on the probability of selecting a fund arises independently of past performance.

Next, I find that the positive effect of previous family ownership on future fund selection obtains even for investors choosing funds with investment objectives that are new to the investor. This finding is consistent with investors assigning their beliefs about the quality or competence of a fund at the family level, while also ruling out two alternative explanations for the repeat buyer results. First, the fact that investors show a preference to remain within the same family even when purchasing a fund in a new objective category shows that my results cannot simply be explained by a clientele effect, whereby individuals cluster in families who specialize in fund objectives that match the preferred investment objective of the investor. Second, this result indicates that the repeat buyer effect does not arise from familiarity resulting from families' tendencies to list member funds from the same objective class in a common prospectus. Finally, consistent with the Mailath and Samuelson (2001) result that reputational beliefs dissipate slowly, I document that individuals' beliefs about funds belonging to older or larger families change slowly, as evidenced by decreased flow-performance sensitivity for these funds.

REFERENCES

Alexander, Gordon, Jonathan Jones, and Peter Nigro, 1998, Mutual fund shareholders: Characteristics, investor knowledge, and sources of information, *Financial Services Review,* 7, 301-316.

Allen, Franklin, 1984, Reputation and product quality, *Rand Journal of Economics*, Autumn, 311-327.

Barber, B. M. and T. Odean, 2000, Trading is hazardous to your wealth: The common stock performance of individual investors, *Journal of Finance*, 55, 773-806.

Barber, B. M. and T. Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics*, 116, 261-292.

Barber, B. M. and T. Odean, 2002, Online investors: Do the slow die first? *Review of Financial Studies*, 15, 455-489.

Barber, B. M. and T. Odean, 2006, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies*, forthcoming.

Barber, Brad M., Terrance Odean, and Lu Zheng, 2005, Out of sight, out of mind: The effects of expenses on mutual fund flows, *Journal of Business,* 78, 2095-2119.

Barberis, N., A. Shleifer and R. Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics*, 49, 307-343.

Battalio, Robert, Andrew Ellul, and Robert Jennings, 2007, Reputation effects in trading on the New York Stock Exchange, *Journal of Finance*, 62 (3), 1243–1271.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy,* 112, 1269–95.

Blume, L., D. Easley, and M. O'Hara, 1994, Market statistics and technical analysis: The role of volume, *Journal of Finance*, 49, 153-181.

Campbell, J., S. Grossman, and J. Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics*, 108, 905-940.

Capon, N., G. J. Fitzsimmons, and R. A. Prince, 1996, An individual level analysis of the mutual fund investment decision, *Journal of Financial Services Research,* 10, 59–82.

Carter, Richard, Frederick H. Dark and Ajai K. Singh, 1998, Underwriter reputation, initial returns, and the long-run performance of IPO stocks, *Journal of Finance*, 53, 285-311.

Carter, Richard and Steven Manaster, 1990, Initial public offerings and underwriter reputation, *Journal of Finance*, 45, 1045–1064.

Chevalier, Judith and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy*, 105, 1167-1200.

Conrad, J. S., A. Hameed, and K. Niden, 1994, Volume and autocovariances in short-horizon individual security returns, *Journal of Finance*, 49, 1305-1329.

De Bondt, W. and R. H. Thaler, 1985, Does the stock market overreact? *Journal of Finance*, 40, 793-805.

Del Guercio, Diane and Paula Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds versus pension funds, *Journal of Financial and Quantitative Analysis*, 37, 523-558.

Diamond, Douglas, 1989, Reputation acquisition in debt markets, *Journal of Political Economy*, 97(4), 828-862.

Diamond, Douglas, 1991, Monitoring and reputation: The choice between bank loans and directly placed debt, *Journal of Political Economy*, 99(4), 689-721.

Elton, Edwin, Martin Gruber and Christopher Blake, 2006, The adequacy of investment choices offered by 401K plans, *Journal of Public Economics*, 90, 1299-1314.

Elton, Edwin, Martin Gruber and Clifton Green, 2007, The impact of mutual fund family membership on investor risk, *Journal of Financial and Quantitative Analysis*, 42, 257-278.

Fama, Eugene, and James D. Macbeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy*, 81, 607 - 636.

Fang, Lily H., 2005, Investment bank reputation and the price and quality of underwriting services, *Journal of Finance*, 60(6), 2729-61.

Fang, Lily and Ayako Yasuda, 2006, Are stars' opinions worth more? The relation between analyst reputation and recommendation values, working paper, Wharton School.

Fang, Lily and Ayako Yasuda, 2007, The effectiveness of reputation as a disciplinary mechanism sell-side research, working paper, Wharton School.

Frieder, Laura and A. Subrahmanyam, 2005, Brand perceptions and the market for common stock, *Journal of Financial and Quantitative Analysis*, 40, 57-85.

Gallaher, Steven, Ron Kaniel and Laura Starks, 2006, Madison Avenue meets Wall Street: Mutual fund families, competition and advertising, working paper, University of Texas at Austin.

George, T. J. and C. Hwang, 2004, The 52-week high and momentum investing, *Journal of Finance*, 59, 2145-2176.

Gervais, S., R. Kaniel, and D. H. Mingelgrin, 2001, The high-volume return premium, *Journal of Finance*, 56, 877-919.

Goetzmann, W. and N. Peles, 1997, Cognitive dissonance and mutual fund investors, *Journal of Financial Research*, 20, 145-158.

Graham, John R. and Alok Kumar, 2006, Do dividend clienteles exist? Evidence on dividend preferences of retail investors, *Journal of Finance*, 61(3), 1305-1336.

Hirshleifer, D., 2001, Investor psychology and asset pricing, *Journal of Finance*, 56, 1533-1597.

Hirshleifer, David A., James N. Myers, Linda A. Myers and Siew Hong Teoh, 2003, Do individual investors drive post-earnings announcement drift? Direct evidence from personal trades, AFA 2004 San Diego Meetings.

Hirshleifer, D. and T. Shumway, 2003, Good day sunshine: Stock returns and the weather, *Journal of Finance*, 58, 1009-1032.

Hong, H., T. Lim, and J. C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance*, 55, 265-295.

Huang, Jennifer, Kelsey Wei and Hong Yan, 2006, Participation costs and the sensitivity of fund flows to past performance, *Journal of Finance*, forthcoming.

Huang, Jennifer, Kelsey Wei and Hong Yan, 2007, Volatility of performance and mutual fund flows, working paper, University of Texas at Austin.

Huberman, Gur and Tomer Regev, 2001, Contagious speculation and a cure for cancer: A nonevent that made stock prices soar, *Journal of Finance*, 56 (1), 387-396.

Investment Company Institute, 1991, *Mutual Fund Fact Book: Industry Trends and Statistics for 1990* (Washington, D.C.).

Investment Company Institute, 1997, Understanding shareholders' use of information and advisers, Washington, DC.

Investment Company Institute, 2006, *Investment Company Fact Book* (Washington, D.C.).

Ippolioto, Richard A., 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law & Economics*, 35, 45-70.

Ivkovic, Zoran and Scott Weisbenner, 2006, Old money matters: The sensitivity of mutual fund redemption decisions to past performance, working paper, University of Illinois.

Ivkovich, Zoran, Sialm, Clemens and Weisbenner, Scott J., 2006, Portfolio concentration and the performance of individual investors, *Journal of Financial and Quantitative Analysis*, forthcoming.

Jackson, Andrew, R., 2005, Trade generation, reputation, and sell-side analysts, *Journal of Finance*, 60, 673-717.

Johnson, Woodrow, 2004, Predictable investment horizons and wealth transfers among mutual fund shareholders, *Journal of Finance*, 59(5), 1979-2012.

Jones, Michael A., Vance P. Lesseig and Thomas I. Smythe, 2005, Financial advisors and multiple share class mutual funds, *Financial Services Review*, 14(1), 1-20.

Khorana, Ajay and Henri Servaes, 2005, An examination of competition in the mutual fund industry, working paper, Georgia Institute of Technology and London Business School.

Klein, Benjamin and Keith B. Leffler, 1981, The role of market forces in assuring contractual performance, *Journal of Political Economy*, 89, 615-641.

Lamont, Owen and Andrea Frazzini, 2007, The earnings announcement premium and trading volume, working paper, Yale School of Management.

Lease, R. C., R. W. Masulis, and J. R. Page, 1991, An investigation of market microstructure impacts on event study returns, *Journal of Finance*, 46, 1523-1536.

Lee, C. M.C., 1992, Earnings news and small traders, *Journal of Accounting and Economics*, 15, 265-302.

Ljungqvist, A., F. Marston, L. Starks, K. Wei, and H. Yan, 2007, Conflicts of interest in sell-side research and the moderating role of institutional investors, *Journal of Financial Economics*, forthcoming.

Mailath, George J. and Larry Samuelson, 2001, Who wants a good reputation? *Review of Economic Studies*, 68, 415-441.

Massa, Massimo, 2001, Why so many mutual funds? Mutual fund families, market segmentation and financial performance, working paper, INSEAD.

Massa, Massimo, 2003, How do family strategies affect fund performance? When performance-maximization is not the only game in town, *Journal of Financial Economics*, 67, 249-304.

Merton, Robert C., 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance*, 42 (3), 483-510.

Milgrom, Paul and John Roberts, 1986, Price and advertising signals of product quality, *Journal of Political Economy*, 94(4), 796-821.

Miller, M. H. and K. Rock, 1985, Dividend policy under asymmetric information, *Journal of Finance*, 40, 1031-1051.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004, Family values and the star phenomenon, *Review of Financial Studies*, 17, 667-698.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2005, The ABCs of mutual funds: On the introduction of multiple share classes, working paper, University of Michigan.

Odean, T., 1999, Do investors trade too much?" *American Economic Review*, 89, 1279-1298.

Seasholes, M. S. and G. Wu, 2007, Predictable behavior, profits, and attention, *Journal of Empirical Finance*, forthcoming.

Sirri, Erik and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance*, 53, 1589-1622.

Stickel, Scott E., 1992, Reputation and performance among security analysts, *Journal of Finance*, 47, 1811-1836.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance*, 62, 1139-1168.

Tirole, Jean, 1996, A theory of collective reputations (with applications to the persistence of corruption and to firm quality), *Review of Economic Studies*, 63, 1-22.

Wall Street Journal, 2003, Putnam investments saw $2 billion outflow in the latest week, December 2, 2003, p. c9.

Zheng, Lu, 1999, Is money smart? A study of mutual fund investors' fund selection ability, *Journal of Finance*, 54, 901-933.

VITA

Michael Charles Yates was born in Birmingham, Alabama on May 5, 1979, the son of Charles Wesley Yates and Marsha Ann Yates. After completing his work at Pelham High School, Pelham, Alabama, in 1997, he entered the University of Alabama in Tuscaloosa, Alabama. He received the degree of Bachelor of Science from the University of Alabama in May 2001 and the degree of Master of Arts from the University of Alabama in May 2002. In September 2002 he entered the Graduate School of The University of Texas at Austin. In September 2007, he will join the faculty of Auburn University as Assistant Professor of Finance.

Permanent Address: 148 Carriagehouse Lane, Auburn, AL 36832

This dissertation was typed by the author.